Optimal Dynamic Information Acquisition*

Weijie Zhong†
January 11, 2022

Abstract. I study a dynamic model in which a decision-maker (DM) acquires information about the payoffs of different alternatives prior to making a decision. The model's key feature is the flexibility of information: the DM can choose any dynamic signal process as an information source, subject to a flow cost that depends on the informativeness of the signal. Under the optimal policy, the DM acquires a signal that arrives according to a *Poisson process*. The optimal Poisson signal confirms the DM's prior belief and is sufficiently precise to warrant immediate action. Over time, given the absence of the arrival of a Poisson signal, the DM continues seeking an increasingly precise but less frequent Poisson signal.

Keywords: dynamic information acquisition, rational inattention, Poisson process
JEL classification: D11, D81, D83

1 Introduction

A decision-maker (DM) often has imperfect information about the payoffs of different alternatives. Therefore, the DM would like to acquire information to learn about the payoffs prior to making a decision. For example, when comparing new technologies, a firm may not know their relative benefits. The firm often spends considerable funds and time on R&D to identify the best technology to adopt. One practically important feature of the information acquisition process is that choosing "what to learn" often involves considering a rich set of salient aspects. In the previous example, when designing the R&D process, the firm decides on aspects such as which technology to test, how much data to collect and analyze, and how intensive the testing should be. Other examples include investors designing algorithms to learn about the returns of different assets and scientists designing experiments to investigate the validity of different hypotheses.

*I am grateful to Sylvain Chassang, Yeon-Koo Che, Mark Dean, Johannes Hörner, Navin Kartik, Qingmin Liu, Konrad Mierendorff, Xiaosheng Mu, Pietro Ortoleva, Andrea Prat, Jakub Steiner, Philipp Strack, Tomasz Strzalecki, Andrzej Skrzypacz, Ming Yang, and participants at conferences and seminars for helpful comments and discussions. This paper is based on the first chapter of Weijie Zhong's Ph.D. dissertation at Columbia University.

†*Graduate School of Business, Stanford University*, Email: `weijie.zhong@stanford.edu`.

To capture such richness, I consider a DM who can flexibly choose "what to learn" and "when to stop learning." The study aims to obtain insight into dynamic information acquisition without restrictions on what type of information can be acquired. In the model, the DM chooses from a set of actions, the payoffs of which depend on a state unknown to the DM. The state is initially selected by nature and remains fixed over time. At any instant in time, the DM chooses whether to stop learning and select an action or to continue learning by flexibly choosing the evolution of the belief process. I introduce two main economic assumptions. (i) The DM discounts delayed payoffs. (ii) Learning incurs a flow cost, which depends convexly on how fast the uncertainty about the unknown state is decreasing. The main model is formulated as a stochastic control-stopping problem in continuous time.

The optimal belief process is a compensated Poisson process. In other words, it is optimal to acquire a *Poisson signal*—a rare and substantial breakthrough that causes a jump in belief. The absence of the breakthrough causes the compensating drift in belief, as per Bayes' rule. The Poisson signal is characterized by three parameters: the *direction*, *magnitude*, and *arrival rate* of the jump. They represent three respective key aspects of learning: the *direction*, *precision*, and *frequency*.

- ***Direction***: The optimal direction of learning is *confirmatory*– the arrival of the Poisson signal induces the belief to jump toward the state the DM currently finds most likely. Conditional on not receiving the signal, the DM gradually becomes less certain about the state.

- ***Precision***: The optimal signal precision is *negatively related* to the continuation value. Therefore, when the DM is less certain about the state, the corresponding continuation value is lower, inducing the DM to seek a more precise Poisson signal.

- ***Frequency***: The optimal signal frequency is *positively related* to the continuation value. In contrast to precision, the optimal signal frequency decreases when the DM is less certain.

- ***Stopping time***: The optimal time to stop learning is immediately after the arrival of the Poisson signal. Therefore, the breakthrough occurs only once at the optimum.

The optimal strategy can be easy to implement in practice. In the context of the R&D process, the optimal strategy involves testing the most promising technology. The optimal test is designed to be challenging to pass; thus, good news is infrequent, as described by a Poisson process. A successful test induces the firm to adopt the tested technology immediately. Otherwise, the firm becomes more pessimistic about the technology and might change its choice of the most promising technology accordingly. Future tests involve higher passing thresholds and lower passing frequencies.

The intuition behind the optimal strategy comes from studying a novel *precision-frequency tradeoff*. Consider a thought experiment of choosing an optimal Poisson signal with a fixed direction and cost level. The remaining two parameters—precision and frequency—are pinned down by the marginal rate of substitution between them. Importantly, the tradeoff depends on the continuation value. Due to discounting, when the continuation value is higher, the DM loses more by delaying the decision. Therefore, it is optimal to acquire a signal more frequently at the cost of lowering the precision to avoid costly delays. That is, the marginal rate of substitution of frequency for precision increases with increasing continuation value. Thus, frequency (precision) is positively (negatively) related to the continuation value.

In addition to precision and frequency, this intuition also explains other aspects. First, given any learning direction, the absence of a signal pushes beliefs away from the target direction. Hence, a signal leading to the same decision quality is more precise and arrives less frequently over time. This dynamic is consistent with the precision-frequency tradeoff when the continuation value decreases over time; that is, the direction of learning is confirmatory. Second, an alternative Gaussian signal leading to belief diffusion is almost surely dominated. The Gaussian signal can be viewed as the limit of a Poisson signal with near-zero precision and infinite frequency. The previous intuition immediately implies that infinite frequency is generally suboptimal except when the continuation value is so high that the DM would like to sacrifice almost all signal precision, which is possible only at the stopping boundaries.

An extensive discussion on the assumptions made in the main model is presented in Section 5. The first key assumption is that the flow cost of information is a convex function of the speed of uncertainty reduction. I show that the speed of uncertainty reduction is the only measure of information that is "path independent" in the sense that the average measure of information acquired per unit time depends only on the overall acquired information and the expected delay but not on the detailed dynamics of the acquisition strategy. This result illustrates that the cost of information does not dictate how information should be chosen dynamically. Hence, the predictions from the main model are implications of how the value of information evolves dynamically in a decision problem. The assumption is then relaxed to accommodate an information cost invariant to the prior belief of the DM. A prior invariant cost, in general, violates the path independence property; hence, the specification of the cost structure has strong implications for the dynamic learning strategy. I illustrate that the Poisson signals remain optimal under certain prior invariant information costs, while the optimal signal process may become approximately Gaussian under others.

Another key assumption is the exponential discounting assumption. Further examining the time preference reveals a deeper connection between dynamic information acquisition and the DM's attitude toward risk in the time dimension: exponential discounting defines a strictly convex utility of time; hence, the DM is time-risk loving. Meanwhile, the optimal Poisson learning strategy involves high time risk: by confirming the more likely state, the DM essentially maximizes the probability of an early stop at the cost of potential long delays. The Gaussian learning strategy, on the other hand, induces a much less dispersed stopping time, thereby making it suboptimal.

The remainder of this paper is structured as follows. Section 2 reviews the related literature. Section 3 introduces the main model. Section 4 characterizes the optimal strategy and illustrates the intuitions. Section 5 presents a discussion of the key assumptions and their extensions. Section 6 concludes and summarizes directions for future research. The key proofs are provided in the appendix. All the remaining proofs are relegated to the online appendix.

2 Related literature

2.1 Dynamic information acquisition

The earliest works on dynamic information acquisition focus on the duration of learning. Wald (1947) and Arrow, Blackwell, and Girshick (1949) analyze a stopping problem where the DM controls the decision time and action choice given exogenous information. Moscarini and Smith (2001) extend the Wald model by allowing the DM to control the precision of a Gaussian signal. A similar Gaussian learning framework is used as the learning-theoretic foundation for the drift-diffusion model by Fudenberg, Strack, and Strzalecki (2018). Following a different route, Che and Mierendorff (2019), Mayskaya (2020), and Liang, Mu, and Syrgkanis (2018) study the sequential choice between multiple exogenously given information sources.[1] The novelty of my framework is that the DM flexibly designs the information generating process. Two concurrent studies, Steiner, Stewart, and Matějka (2017) and Hébert and Woodford (2018), also model flexible dynamic information acquisition; however, the incentive to optimally choose what to learn over time is absent. In Steiner, Stewart, and Matějka (2017), the linear flow cost assumption makes instantaneous learning optimal, whereas

[1]The sequential search models (Callander (2011), Doval (2018), Ke and Villas-Boas (2019), Klabjan, Olszewski, and Wolinsky (2014), and Weitzman (1979)) and multi-armed bandit models (Bergemann and Välimäki (1996), Bolton and Harris (1999), Gittins (1974), and Weber et al. (1992)) also feature dynamic information acquisition. In these models, the DM has control over whether to reveal an option or pull an arm sequentially. However, the information provided by the actions is exogenously given.

the no-discounting assumption makes all dynamic learning strategies payoff equivalent in Hébert and Woodford (2018).[2]

2.2 Rational inattention

This study is a dynamic extension of static rational inattention (RI) models. The entropy-based RI framework is first introduced in Sims (2003). Matějka and McKay (2014) study the flexible information acquisition problem with mutual information as cost and justify a generalized logit decision rule.

I employ the speed of uncertainty reduction to measure the amount of information acquired per unit time. This measure features the uniform posterior separability (UPS) property introduced by Caplin and Dean (2013). The property generalizes mutual information and is widely used in applications (Clark and Reggiani (2021), Gentzkow and Kamenica (2014), Matyskova (2018), and Rappoport and Somma (2017)). Several foundations have been provided for UPS (Caplin, Dean, and Leahy (2017), Frankel and Kamenica (2019), and Morris and Strack (2019)). Section 5 presents a novel implication of UPS in dynamic settings.

2.3 Information design

This study employs a belief-based approach to model the choice of information. This approach is widely used for studying Bayesian persuasion models (Ely (2017), Kamenica and Gentzkow (2011), and Mathevet, Perego, and Taneva (2019 forthcoming)). An important methodology in this literature is the concavification method developed in Aumann, Maschler, and Stearns (1995) (based on Carathéodory's theorem). I generalize the concavification method to solve an auxiliary discrete-time problem.

3 Model setup

The main model is a continuous-time stochastic control problem. A DM chooses an irreversible action at an endogenous decision time. The DM can flexibly control the information received before the decision time, bearing a cost on the information.

Decision problem: Time $t \in [0, +\infty)$. The DM discounts the delayed utility at rate $\rho > 0$. The DM is a von Neumann-Morgenstern expected utility maximizer with Bernoulli utility associated with action-state pair $(a, x) \in A \times X$ at time t given by $e^{-\rho t} u(a, x)$. Both the action space A and the state space X are finite. The DM holds a prior belief $\mu \in \Delta(X)$ about the state. Define $F(\nu) \triangleq \max_{a \in A} E_\nu[u(a, x)]$ for all beliefs $\nu \in \Delta(X)$.

Information: I model information using a belief-based approach. The DM chooses

[2] Steiner, Stewart, and Matějka (2017) study a more general repeated decision-making problem that is history-dependent, which generates nontrivial dynamics in the signal process.

the entire posterior belief process $\langle \mu_t \rangle$ in a nonparametric way. Bayes' rule is required to be satisfied at every instant of time — $\forall s > t$, $\mathbb{E}[\mu_s | \mathcal{F}_t] = \mu_t$. Thus, I restrict $\langle \mu_t \rangle$ to be a martingale, with $\langle \mathcal{F}_t \rangle$ being its natural filtration.[3]

Cost of information: Given a belief process $\langle \mu_t \rangle$, I assume that the DM pays a flow cost that depends on the "amount" of information acquired per unit of time. The flow cost is $C(I_t)$, where $C : \mathbb{R}^+ \to \mathbb{R}^+$ is the cost function and I_t is the *flow information measure* that quantifies the informativeness of signals.

To define the flow information measure, it is useful to introduce the following definitions. $\forall \langle \mu_t \rangle$, $f \in C(\Delta(X))$ and $s > 0$, define a process $\mathcal{A}_t^s f \triangleq \mathbb{E}\left[\frac{f(\mu_{t+s}) - f(\mu_t)}{s} \middle| \mathcal{F}_t \right]$. Let $\mathscr{D}(f)$ be the domain of $\langle \mu_t \rangle$ such that when $s \to 0$, $\langle \mathcal{A}_t^s f \rangle$ converges in probability to a bounded process. Denote the limit by $\langle \mathcal{A}_t f \rangle$. $\langle \mathcal{A}_t f \rangle$ captures the "speed" at which $\langle f(\mu_t) \rangle$ changes at every instant of time.[4]

Assumption 1. *$I_t = -\mathcal{A}_t H$, where $H \in C^2 \Delta(X)$ is strictly concave.*[5]

Given concave H, $-\mathcal{A}_t^s H = \mathbb{E}\left[\frac{H(\mu_t) - H(\mu_{t+s})}{s} \middle| \mathcal{F}_t \right]$ increases in the mean-preserving spread order of the conditional distribution of μ_{t+s}. Therefore, the information measures that satisfy Assumption 1 are compatible with the Blackwell order on signal informativeness. One can interpret function H as a measure of uncertainty (introduced by Frankel and Kamenica (2019)). Thus, I_t is the speed at which uncertainty falls when belief updates. Assumption 1 nests standard notions such as mutual information rate when H is Shannon's entropy (introduced by Shannon (1948)) and quadratic variation rate when H is quadratic. In Section 5, I discuss the implications of Assumption 1 in detail and present extensions.

Assumption 2. *$C : \mathbb{R}^+ \to \mathbb{R}^+$ is weakly increasing and convex. $\lim_{I \to \infty} C'(I) = \infty$.*

Assumption 2 states that the marginal cost of information is positive, increasing, and unbounded. An important implication of Assumption 2 is that the DM has strict incentive to smooth the cost of information over time. In Section 5.2, I analyze an alternative setting where the marginal cost of information is constant.

Stochastic control: The DM solves the following stochastic control problem:

$$V(\mu) = \sup_{\langle \mu_t \rangle \in \mathbb{M}, \tau} E\left[e^{-\rho\tau} F(\mu_\tau) - \int_0^\tau e^{-\rho t} C(I_t) \mathrm{d}t \right], \tag{1}$$

[3] This approach is the continuous-time analogy of the belief martingale model in Ely, Frankel, and Kamenica (2015).

[4] As a special case, when $f \in C(\Delta(X))$, $\mathscr{D}(f)$ contains all Feller processes, on which $\mathcal{A}f$ reduces to the standard infinitesimal generator (subscript t omitted due to the Markov property of Feller processes).

[5] Formally H is C^2 when restricted to the interior of $\Delta(X')$ for any $X' \subset X$.

where $\mathbb{M}$ is the set of all cadlag martingales $\langle \mu_t \rangle$ in $\mathscr{D}(H)$ satisfying $\mu_0 = \mu$, and τ is a $\langle \mathcal{F}_t \rangle$-measurable stopping time. The technical discussion on the integrability in Equation (1) is relegated to Remark A.1 in Appendix A.

The objective function in Equation (1) is straightforward. The DM acquires information that affects $\langle \mu_t \rangle$ and chooses the stopping time τ to maximize the expected stopping payoff $E[e^{-\rho\tau}F(\mu_\tau)]$ less the total information cost $E\left[\int_0^\tau e^{-\rho t}C(I_t)\mathrm{d}t\right]$. The novel feature here is that the DM can fully control $\langle \mu_t \rangle$.

3.1 *Dynamic programming and Hamilton-Jacobi-Bellman equation*

It is useful to apply the principle of dynamic programming to obtain the Hamilton-Jacobi-Bellman (HJB) equation of the stochastic control problem:

$$\max\Big\{\underbrace{F(\mu_t)-V(\mu_t)}_{\text{stopping value}},\ \underbrace{-\rho V(\mu_t)}_{\text{discount}}+\sup_{\mathrm{d}\mu_t}\Big\{\underbrace{\mathcal{A}_t V}_{\text{continuation value}}-\underbrace{C(-\mathcal{A}_t H)}_{\text{control cost}}\Big\}\Big\}=0, \tag{2}$$

where $\mathcal{A}_t V$ is the flow utility gain from continuing. $\mathcal{A}_t V$ and $\mathcal{A}_t H$ are determined by the belief process in the neighborhood of t, denoted by $\mathrm{d}\mu_t$. At any instant in time when the control is chosen optimally, either stopping is optimal (the first term is 0) or continuing is optimal and the net continuation gain equals the loss from discounting (the second term is 0). Utilizing the HJB equation to solve Equation (1) requires verification and simplifying the abstract operator $\mathcal{A}_t$. The existing theories on stochastic control have little power for both tasks.[6] In Theorem 1, I achieve both goals by showing that the solution of Equation (1) is characterized by a simple parametric HJB equation.

Theorem 1. *Given Assumptions 1 and 2, if $V(\mu) \in C^1\Delta(X)$ solves*

$$\rho V(\mu)=\max\Big\{\rho F(\mu),\sup_{p,\nu,\sigma} p(V(\nu)-V(\mu)-\nabla V(\mu)(\nu-\mu))+\frac{1}{2}\sigma^T \mathrm{Hess}V(\mu)\sigma \\ -C\Big(p(H(\mu)-H(\nu)+\nabla H(\mu)(\nu-\mu))-\frac{1}{2}\sigma^T\mathrm{Hess}H(\mu)\sigma\Big)\Big\}, \tag{3}$$

where $(p,\nu,\sigma) \in \mathbb{R}^+ \times \Delta(\mathrm{Supp}(\mu)) \times \mathbb{R}^{|\mathrm{Supp}(\mu)-1|}$, then $V(\mu)$ solves Equation (1).[7]

Proof. See Appendix A. □

---

[6] The existing verification results based on the martingale methods do not nest Equation (1) (Boel and Kohlmann (1980), Davis (1979), and Striebel (1984)). Moreover, the martingale methods (e.g. theorem 4.3.1 of Boel and Kohlmann (1980)) does not provide an explicit representation of the abstract $\mathcal{A}_t$ operator, which is considered their main drawback (see the discussions in Davis (1979)).

[7] $V(\mu) \in C^1\Delta(X)$ solves Equation (3) in terms of viscosity solutions. Formally, the notation $\sigma^T\mathrm{Hess}V(\mu)\sigma$ in Equation (3) denotes $D^2V(\mu,\sigma)||\sigma||^2$, where $D^2V(\mu,\sigma)=\overline{\lim_{\delta\to 0}}\, 2\frac{V(\mu+\delta\sigma)-V(\mu)-\nabla V(\mu)\delta\sigma}{\delta\|\sigma\|^2}$. See Remark A.2 for detailed discussions.

In Equation (3), ∇ and Hess are gradient and Hessian operators, respectively. Theorem 1 first states that $V(\mu)$ is characterized by an HJB equation. Moreover, the abstract HJB Equation (2) can be simplified to a parametric equation (3). As a direct corollary, the stochastic control problem Equation (1) can be solved by considering a simple family of jump-diffusion processes:

$$\mathrm{d}\mu_t = \underbrace{(\nu(\mu_t) - \mu_t)(\mathrm{d}J_t(p(\mu_t)) - p(\mu_t)\mathrm{d}t)}_{\text{compensated Poisson process}} + \underbrace{\sigma(\mu_t)\mathrm{d}W_t}_{\text{Gaussian diffusion}},$$

where $(p, \nu, \sigma) : \mu_t \mapsto \mathbb{R}^+ \times \Delta(\mathrm{Supp}(\mu)) \times \mathbb{R}^{|\mathrm{Supp}(\mu)|-1}$ are the control parameters, $J_t(\cdot)$ is a Poisson counting process with Poisson rate $(\cdot)$, and W_t is a standard one-dimensional Wiener process.[8] This is because Itô's lemma implies an explicit representation of the infinitesimal generator for this family:

$$\mathcal{A}V(\mu) = \underbrace{p(\mu)(V(\nu(\mu)) - V(\mu) - \nabla V(\mu)(\nu(\mu) - \mu))}_{\text{flow value of Poisson jump \& drift}} + \underbrace{\frac{1}{2}\sigma(\mu)^T \mathrm{Hess} V(\mu)\sigma(\mu)}_{\text{flow value of diffusion}}.$$

The compensated Poisson jump and Gaussian diffusion parts in the stochastic differential equation represent two different types of learning strategies.

- ***Poisson learning***: The DM uses *Poisson learning* or acquires a *Poisson signal* when a compensated Poisson part exists in the belief process. A Poisson jump in the belief process can be induced by observing a rare breakthrough, the arrival of which follows a Poisson process. The compensating belief drift is induced by observing no breakthroughs. The control variables for Poisson learning are (p, ν), representing three endogenously relevant aspects of Poisson learning. The arrival rate p represents the *frequency* of learning. ν is the destination of the belief jump. The jump direction represents the *direction* of learning. The jump magnitude represents the *precision* of learning.

- ***Gaussian learning***: The DM uses *Gaussian learning* or acquires a *Gaussian signal* when a diffusion part exists in the belief process. Gaussian diffusion in the belief process can be induced by observing the realization of a Gaussian process whose unobservable drift is the state x. The flow variance σ represents the signal precision.

[8] The reduction to jump-diffusion processes is similar in form to the Lévy-Itô decomposition and hence may be expected. Theorem 1 differs from existing decomposition theorems in that it applies to more general processes but only applies at the optimum.

Equation (3) suggests that the DM considers four types of tradeoffs to determine the optimal strategy: (i) the standard continuing-stopping tradeoff in optimal stopping problems, captured by the outer-layer maximization; (ii) the information cost-utility gain tradeoff, which determines the total cost spent on learning; (iii) the Poisson-Gaussian tradeoff, which determines the proportion of cost allocated to the Poisson and the Gaussian signals; and (iv) the precision-frequency tradeoff, which determines the marginal rate of substitution of signal frequency for precision. These tradeoffs, especially the precision-frequency tradeoff, are discussed in detail to characterize the solution of Equation (3) in Section 4.[9]

4 Optimal information acquisition

In this section, I prove that Equation (3) admits a solution and characterize the solution, assuming a binary state. Then, in Section 4.2, I discuss the key tradeoffs and provide intuition for the optimal strategy.

4.1 Main characterization theorem

Assumption 3.

(i) Binary state: $|X| = 2$.

(ii) Positive payoff: $\forall \mu \in [0,1],\ F(\mu) > 0$.

(iii) Uncertainty measure: $H''(\mu) < 0$ *and is locally Lipschitz on* $(0,1)$, $\lim_{\mu\to 0,1} |H'(\mu)| = \infty$.

(iv) Cost function: $C \in C^2(\mathbb{R}^+)$, $C(0) = 0$, $C'(0) \geq 0$, *and* $C''(I) > 0$.

Assumption 3 restricts the state space to be binary and imposes technical conditions on F, H, and C for tractability. First, the binary state assumption implies that the belief space is one-dimensional. Thus, ordinary differential equation (ODE) theory can be applied to construct a candidate solution. Although the existence of the solution relies on the binary state assumption, Section 5.3 shows that the characterization generalizes to more than two states. Second, the utility from decision-making is assumed to be strictly positive. Third, H has locally Lipschitz continuous second derivative, is strictly convex, and satisfies an Inada condition.[10] Forth, the cost function $C(I)$ is twice continuously differentiable, increasing, strictly convex, and $C(0) = 0$.

Theorem 2. *Given Assumption 3, there exists a quasi-convex value function* $V \in C^1[0,1]$ *solving Equation (3). Let* $E = \{\mu \in [0,1] | V(\mu) > F(\mu)\}$. $\exists$ *a.e. unique policy functions*

[9]The precision-frequency tradeoff nests the Poisson-Gaussian tradeoff as a limiting case, and can be used to show that Gaussian learning is almost surely suboptimal under further technical assumptions.

[10]The condition $\lim_{\mu\to 0,1} |H'(\mu)| = \infty$ is imposed to guarantee non-degenerate stopping regions. The condition only restricts the boundary behavior of H. It allows for Shannon's entropy but rules out functions with bounded derivatives (e.g. quadratic variation).

$(p,\nu): E \to \mathbb{R}^+ \times [0,1]$ *satisfying*

$$\rho V(\mu) = p(\mu)\big(F(\nu(\mu)) - V(\mu) - V'(\mu)(\nu(\mu)-\mu)\big) \\ - C\big(p(\mu)\big(H(\mu) - H(\nu(\mu)) + H'(\mu)(\nu(\mu)-\mu)\big)\big).$$

There exists $\mu^* \in \arg\min V$ *s.t.*

(i) *Poisson learning:* $\rho V(\mu) > \max_\sigma \frac{1}{2}\sigma^2 V''(\mu) - C\left(-\frac{1}{2}\sigma^2 H''(\mu)\right)$, $\forall \mu \in E \setminus \mu^*$.

(ii) *Direction:* $\mu > (<) \mu^* \implies \nu(\mu) > (<) \mu$.

(iii) *Precision:* $\nu(\mu)$ *decreases in each connected region of* $E \setminus \mu^*$.

(iv) *Intensity:* $I(\mu) \triangleq p(\mu)(H(\nu(\mu)) - H(\mu) - H'(\mu)(\nu(\mu)-\mu))$ *is increasing in* $V(\mu)$.

(v) *Stopping time:* $\nu(\mu) \in E^C$.

Theorem 2 proves that Equation (3) admits a solution and characterizes the optimal policy function. The theorem first states that the optimal value function is implemented by a *Poisson signal*; that is, seeking a breakthrough that causes the belief to jump to $\nu(\mu)$ at rate $p(\nu)$. It follows immediately from Theorem 1 that under Assumptions 1 and 3 the stochastic control problem Equation (1) is solved by the compensated Poisson process:

$$\mathrm{d}\mu_t = (\nu(\mu_t) - \mu_t)(\mathrm{d}J_t(p(\mu_t)) - p(\mu_t)\mathrm{d}t).$$

Theorem 2 then characterize the optimal policy. Property (i) states that the payoff from acquiring a Gaussian signal (the RHS of the inequality) is strictly suboptimal, except for at most one critical belief. The optimal Poisson signal is characterized by four aspects of learning and the stopping time.

Direction: Property (ii) states that the optimal direction is *confirmatory*: when $\mu > \mu^*$, the DM holds a high prior belief (relative to μ^*) for state 1 and acquires a signal whose arrival induces an even higher posterior belief $\nu(\mu)$ and vice versa for $\mu < \mu^*$.

Precision: Property (iii) implies that the optimal precision measured by $|\nu(\mu) - \mu^*|$ is *negatively related* to the certainty of the belief (measured by $|\mu - \mu^*|$). Since $\mu^* \in \arg\min V$, the property equivalently states that precision is negatively related to the continuation value.

Intensity: Property (iv) states that the optimal intensity measured by $I(\mu)$ is *positively related* to the continuation value.

Frequency: The previous two properties imply that the optimal frequency $p(\mu)$ is *positively related* to the continuation value.

Stopping time: Property (v) states that the image of ν is in the stopping region. That is, the optimal stopping time is exactly the signal arrival time.

Combining these properties pins down the optimal learning dynamics. The DM seeks a signal that arrives according to a Poisson process. The arrival of the signal confirms the DM's prior belief and is sufficiently accurate to warrant immediate action. Absent the arrival of a Poisson signal, the DM becomes less certain about the state, following Bayes' rule. The DM's continuation value decreases correspondingly; hence, she invests less in learning and seeks a Poisson signal with a lower frequency and higher precision. [11] The DM's belief may reach μ^* before the arrival of the signal. Then, she seeks a mixture of two signals that confirm each state in a balanced way such that her posterior belief conditional on observing no signal stays at μ^*. The learning process is stationary afterwards. [12]

I provide two examples illustrating Theorem 2. Example 1 is a minimal working example with two actions. Example 2 has four actions, and provides richer implications.

Example 1. $F(\mu) = \max\{2\mu - 1, 1 - 2\mu\}$, $H(\mu) = -\mu \log(\mu) - (1-\mu)\log(1-\mu)$ is Shannon's entropy, $\rho = 1$, and $C(I) = \frac{1}{2}I^2$. The solution is presented in Figures 1 and 2. In Figure 1-(a), the dashed lines depict $F(\mu)$, the blue curve depicts $V(\mu)$, and the blue shaded region is the experimentation region E. Figure 1-(b) shows the optimal posterior $\nu(\mu)$ as a function of the prior. The three arrows in Figure 1-(a) start at the priors and point to their optimal posteriors. The blue curve in Figure 1-(c) shows the optimal intensity $I(\mu)$ as a function of the prior. Clearly, $I(\mu)$ is a monotonic transformation of $V(\mu)$ in the experimentation region.

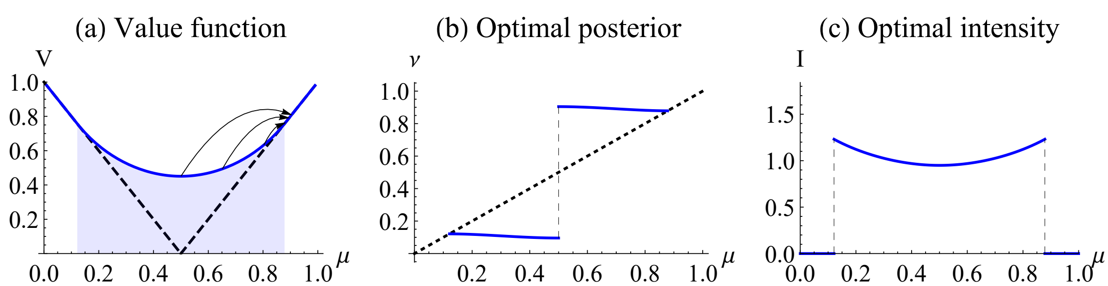

Figure 1: Value and policy functions

Figure 2 illustrates the dynamics of the optimal policy, given $\mu = 0.85$. Figure 2-(a) depicts the optimal belief process. Conditional on no signal arrival, the posterior be-

[11] As per Property (iii), the monotonicity holds in each connected region of $E \setminus \mu^*$ but not globally. However, for a given prior, property (v) implies that the belief process moves continuously within a connected region until it jumps and stops immediately. Hence, the monotonicity holds on all relevant beliefs. See an illustration in Example 2.

[12] Formally, the DM's optimal policy $(\nu(\mu), p(\mu))$ involves only one posterior belief. However, when belief is μ^* she may alternate between the two directions infinitely fast that the belief process is as if she mixes two signals.

lief drifts toward $\mu^* = 0.5$. In this example, two *phases* of learning occur (represented by different colors of shaded regions in Figure 2-(a)). In the first phase (blue region), the DM seeks a Poisson signal to confirm the most likely state. As time passes, signal precision increases while signal frequency and learning intensity decrease (as in Figure 2-(b)&(c)). Eventually, the DM believes the two states are equally likely and switches to the second phase (gray region) where she seeks two signals that confirm each state in a balanced way.

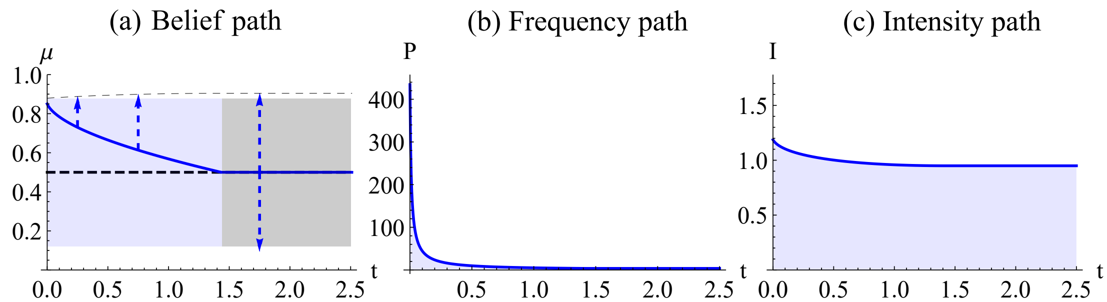

Figure 2: Dynamics of optimal policy

Example 2. There are four actions, whose expected payoffs are represented by the four dashed lines in Figure 3-(a). The two blue lines are steeper, indicating that the corresponding actions involve higher risk in payoffs. The two red lines are flatter, and thus the corresponding actions are safer. The experimentation region contains three disjoint intervals. For the middle interval, in the red regions, the DM has a more certain belief and searches for a signal that confirms a safer action (red arrow). In the blue region, the DM has a more uncertain belief and searches for a riskier action (blue arrow). Figure 3-(c) depicts the optimal belief process with a prior belief in the red region. The experimentation follows three phases. The DM acquires a signal leading to a safer action in the first phase; acquires a signal leading to a riskier action in the second phase; and acquires two signals that confirm each state in a balanced way in the third. The signal precision increases monotonically over time, as represented by the thin dashed curve in Figure 3-(c). In more general settings, the optimal strategies are qualitatively the same as that in Example 2, except for potentially more phases, corresponding to more actions to be considered.

4.2 *Proof methodology and key intuitions*

At the end of Section 3, I introduced four types of tradeoffs. Now, I discuss the tradeoffs in detail and illustrate how they determine the optimal strategy. I first derive

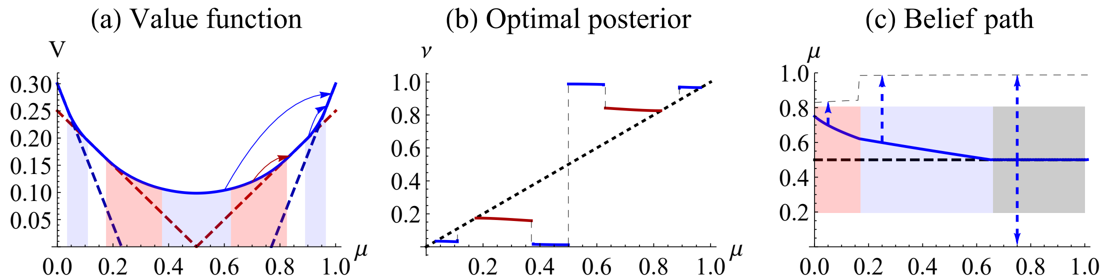

Figure 3: Example with four actions

a geometric characterization that facilitates visualizing the optimal policy. Then, I analyze the key tradeoffs using the characterization and provide intuitions for the optimal policy. Section 4.2.2 outlines a proof for Theorem 2.

4.2.1 *Geometric representation and key tradeoffs*

A thought experiment is useful to gain intuition. Fix the value function V and optimize the RHS of Equation (3) using only Poisson signals:

$$\sup_{p,\nu} p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) - C\big(p\big(H(\mu) - H(\nu) + H'(\mu)(\nu-\mu)\big)\big). \tag{4}$$

Change variables by defining $I = p(H(\mu) - H(\nu) + H'(\mu)(\nu-\mu))$. Thus, Equation (4) becomes

$$\sup_{I,\nu} \frac{V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{H(\mu) - H(\nu) + H'(\mu)(\nu-\mu)} \cdot I - C(I).$$

The problem is separable in I and ν. The maximizer (ν^*, I^*) is characterized by:

$$\begin{cases} \nu^* \in \arg\max_\nu \frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{H(\mu)-H(\nu)+H'(\mu)(\nu-\mu)} \\ C'(I^*) = \max_\nu \frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{H(\mu)-H(\nu)+H'(\mu)(\nu-\mu)}. \end{cases} \tag{5}$$

Define $\lambda = C'(I^*)$ and $G(\mu) = V(\mu) + \lambda H(\mu)$. $G(\mu)$ is termed the *gross value function*. Thus, Equation (5) implies:

$$\begin{aligned} &\begin{cases} \frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{H(\mu)-H(\nu)+H'(\mu)(\nu-\mu)} \le \lambda & ,\forall \nu \in [0,1] \\ \frac{V(\nu^*)-V(\mu)-V'(\mu)(\nu^*-\mu)}{H(\mu)-H(\nu^*)+H'(\mu)(\nu^*-\mu)} = \lambda \end{cases} \\ \Longrightarrow &\begin{cases} G(\nu) \le G(\mu) + G'(\mu)(\nu-\mu) & ,\forall \nu \in [0,1] \\ G(\nu^*) = G(\mu) + G'(\mu)(\nu^*-\mu). \end{cases} \end{aligned} \tag{6}$$

Equation (6) states that $G(\nu)$ is everywhere below the tangent line of G at μ, except that $G(\nu^*)$ touches the tangent line. See Figure 4 for a graphical illustration.[13] Figure 4-(a)

[13] All figures in Section 4.2.1 are created using the parameters in Example 1.

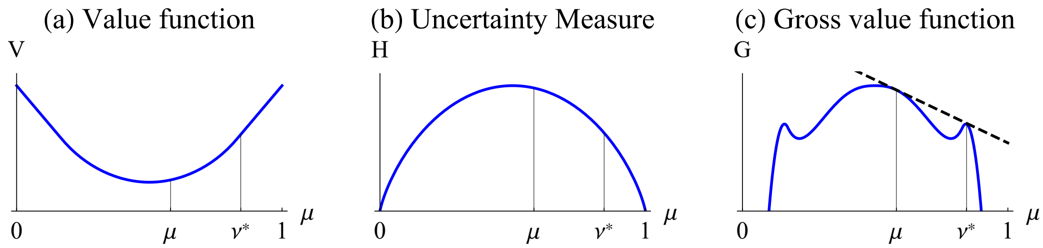

Figure 4: Concavification of the gross value function

and Figure 4-(b) depict the value function V and the uncertainty measure H, respectively. Figure 4-(c) depicts the gross value function $G = V + \lambda H$. $G(\mu)$ and $G(\nu^*)$ share the same tangent line (the dashed line in Figure 4-(c)). When ν^* is unique, μ and ν^* are the two end points of a concavified region (the interval on which $G < \text{co}(G)$). [14]

Equation (6) is an analogy to the concavification method in Bayesian persuasion problems (Aumann, Maschler, and Stearns (1995) and Kamenica and Gentzkow (2011)). The difference is that in a Bayesian persuasion problem, the boundary points of a concavified region are the optimal posteriors, whereas the prior is also on the boundary of a concavified region in the current problem. This property has clear economic meaning. G combines the value function V and the uncertainty measure H using the multiplier $\lambda = C'(I^*)$. Since C' is unbounded by Assumption 3, at optimum, the flow cost must be bounded. Thus, the multiplier λ always adjusts such that the total reduction of uncertainty is in the same order of dt. That is, the shape of G changes endogenously to push the prior belief μ toward one of the optimal posterior belief.

Now, suppose the HJB equation is satisfied; that is, Equation (4) equals the flow discounting loss $\rho V(\mu)$. Then combining Equations (3) and (5) implies:

$$\rho V(\mu) = I^* C'(I^*) - C(I^*) \tag{7}$$

Combining Equation (6) and Equation (7) identifies the value function V and corresponding strategies p, ν. Now, I analyze the key tradeoffs in the dynamic information acquisition problem by studying Equations (6) and (7).

1. Utility gain versus information cost

Equation (7) illustrates the utility gain versus information cost tradeoff. Since C is convex, $IC'(I) - C(I)$ increases in I; that is, the optimal flow informativeness measure I increases in the continuation value. The intuition for this property is discussed in Moscarini and Smith (2001). The marginal gain from experimentation is proportional to the continuation value, while the marginal cost increases in I. Therefore, the optimal

[14] In the formal proof, I verify that except for non-generic cases, ν^* is unique. In this section I invoke this uniqueness and always treat (μ, ν^*) as the end points of a concavified region.

cost increases in the value function. This property is called "value-level monotonicity" in Moscarini and Smith (2001), where the "level" (the flow variance of a diffusion process) is a parameter for both the cost and precision of a Gaussian signal. My analysis differentiates this intuition separately from another important tradeoff between signal precision and frequency. I refer to the property as "value-intensity monotonicity", where the information measure I represents the intensity of learning.

An implication of the value-intensity monotonicity is that $\lambda = C'(I^*)$ increases in the continuation value. Having characterized λ, I can proceed to Equation (6).

2. Precision versus frequency

A novel tradeoff characterized by Equation (6) is the precision versus frequency tradeoff. Given the total intensity I, the DM allocates I to precision and frequency. The following illustrates how this tradeoff changes as the continuation value changes.

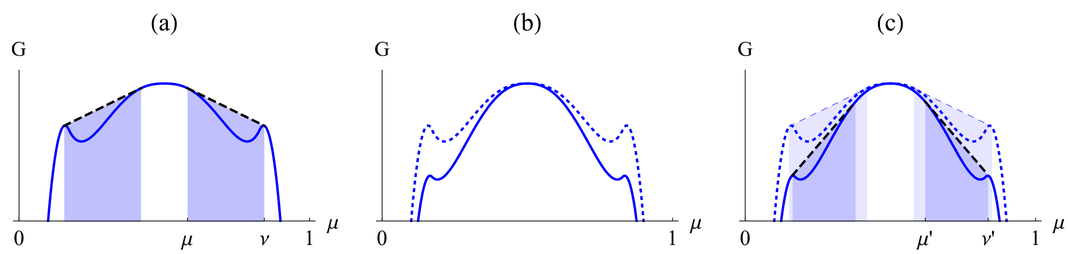

Figure 5: Precision-frequency tradeoff

Figure 5 shows how varying λ affects the optimal jump magnitude. In Figure 5-(a) the blue curve is G, and the dashed curve is $\text{co}(G)$ (G's upper concave hull). I call the blue region, where $G(\mu) < \text{co}(G)(\mu)$, the *concavified region* and the white region, where $G(\mu) = \text{co}(G)(\mu)$, the *concave region*. The prior μ and optimal posterior ν are the end points of a concavified region. Consider $G_1 = V + \lambda_1 H$, where $\lambda_1 > \lambda$. Figure 5-(b) depicts G (the dashed curve) and G_1 (the blue curve). Since G_1 is G plus a strictly concave function, any belief in the concave region of G remains in the concave region of G_1. Thus, as λ increases, the white region expands and the blue region contracts (see Figure 5-(c)). Hence, the prior and optimal posterior move closer together. Recall that λ is monotonic in V; thus, the DM is more willing to choose a signal that induces a shorter belief jump when the continuation value is higher.

The intuition for this property is as follows. When the DM is more certain about the state, the continuation value is higher; hence, the utility loss from discounting is higher. The DM wants to receive a signal more frequently to benefit from the high value sooner. In other words, the marginal rate of substitution of frequency for precision increases in the continuation value.

Confirming versus contradicting: Now, I use the concavification method to pin down the optimal jump direction. Figure 6 zooms into one concavified region $[\mu, \nu]$. The blue

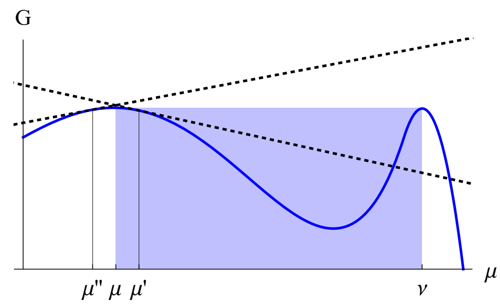

Figure 6: Confirming versus contradicting

curve is $G(\mu)$ and the shaded region is the concavified region. For any $\mu' \in (\mu, \nu)$, $G(\mu') < \text{co}(G)(\mu)$. Therefore, the tangent line of G at μ' must cross G somewhere (the dotted line). This observation implies that for the optimality condition Equation (6) to be satisfied at μ', the gross value function must be strictly more concave than G; hence, the multiplier $\lambda(\mu')$ must be higher than $\lambda(\mu)$. Thus, the value-intensity monotonicity implies that $V(\mu') > V(\mu)$. A similar argument suggests that for $\mu'' \notin (\mu, \nu)$, the multiplier λ must be lower and $V(\mu'') < V(\mu)$. Therefore, the optimal belief jumps toward the direction where the value function increases. That is, the optimal signal is confirmatory.

The analysis is consistent with the intuition. Absent a signal, the posterior drifts away from the target belief. Thus, the learning precision should increase over time to make up for the belief drift. The increasing precision over time is only consistent with a decreasing value function given the precision-frequency tradeoff. Therefore, the optimal posterior is always in the direction where the value function increases.

Poisson versus Gaussian: Thus far, I have ignored the possibility of Gaussian signals. In fact, Gaussian signals are implicitly modeled in Equation (6). Consider the optimization w.r.t. Gaussian signals:

$$
\begin{aligned}
&\sup_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - C\left(-\frac{1}{2}\sigma^2 H''(\mu)\right) \\
\Longrightarrow\ &\text{FOC}: V''(\mu) + \lambda H''(\mu) = 0 \\
\Longleftrightarrow\ &G''(\mu) = 0 \qquad (8)
\end{aligned}
$$

where $\lambda = C'\left(-\frac{1}{2}\sigma^2 H''(\mu)\right)$. The comparison of Equations (6) and (8) shows that Equation (8) is exactly the limit of Equation (6) when the posterior ν converges to the prior μ. This result is intuitive since a Gaussian signal can be approximated as a Poisson signal with very low precision and high arrival rate. Therefore, this tradeoff is a special case of the precision-frequency tradeoff. Selecting a Gaussian signal is a corner solution when the DM wants to sacrifice almost all precision for frequency—a slightly less patient DM is willing to avoid any waiting and stop immediately, while a slightly more patient DM is willing to wait for a more precise Poisson signal. Therefore, the Gaussian signal is optimal only on the boundaries of the experimentation regions.

3. Continuing versus stopping

Theorem 2 states that repeated jumps are suboptimal. I illustrate in Figure 7 based on the concavification method that repeated jumps can be improved by a direct jump. Let ν be the optimal posterior and $G(\cdot) = V(\cdot) + \lambda(\mu)H(\cdot)$ be the gross value function

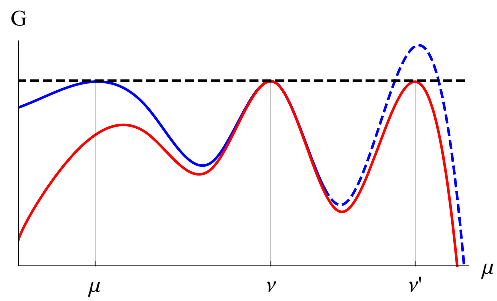

Figure 7: Continuing versus stopping

corresponding to prior μ (solid blue curve). Hypothetically, imagine that it is optimal to continue at ν and the optimal posterior is ν'. As already shown in the analysis of the optimal direction, $V(\nu') > V(\nu) > V(\mu)$. Let $G_1(\cdot) = V(\cdot) + \lambda(\nu)H(\cdot)$ be the gross value function corresponding to prior ν. Then, G_1 is strictly more concave than G, as depicted by the red curve in Figure 7. For a better illustration, I shift G_1 by a linear function such that it shares the same tangent line at ν with G. The optimality condition Equation (6) implies that $G(\mu)$, $G(\nu)$, $G_1(\nu)$, and $G_1(\nu')$ are on the same tangent line. Hence, since G is strictly more convex than G_1, $G(\nu') > G_1(\nu') = G(\mu) + G'(\mu)(\nu' - \mu)$, as depicted by the dashed blue curve in Figure 7. Given the definition of G, $G(\nu') > G(\mu) + G'(\mu)(\nu' - \mu) \iff \frac{V(\nu') - V(\mu) - V'(\mu)(\nu' - \mu)}{H(\mu) - H(\nu') + H'(\mu)(\nu' - \mu)} > \lambda(\mu)$. Therefore, directly jumping to ν' from μ yields strictly higher payoff than jumping to ν.

Example 2 (continued)**.** I continue Example 2 by illustrating the concavification method in its setting. In Figure 8, I plot the gross value functions corresponding to different prior beliefs. In each figure, μ is the prior, ν is the optimal posterior, and the dashed line tangents the gross value function at the prior belief. The colored bands illustrate the experimentation regions. In the blue (red) region, the optimal posterior induces the choice of riskier (safer) actions. In Figure 8-(a), $\mu = 0.5 \in \arg\min V$; hence, G is most convex. Since $\mu = \mu^*$, jumping to either side and choosing a riskier action is optimal. In

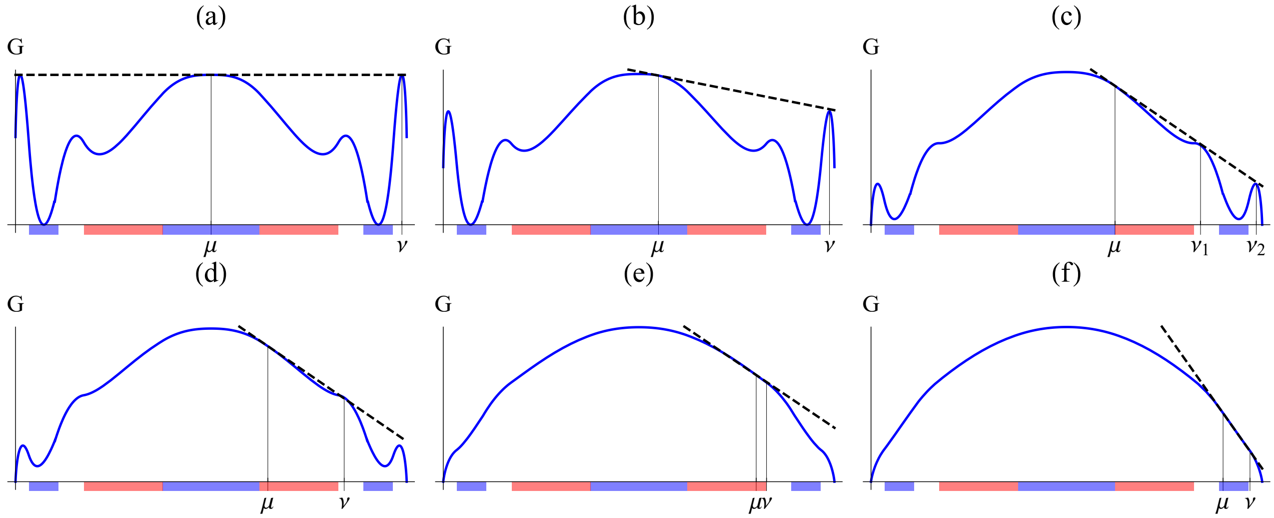

Figure 8: Concavification of the gross value function

Figure 8-(b), μ is slightly higher; hence, G becomes more concave and the optimal precision decreases. In Figure 8-(c), μ is at the threshold between the blue and red regions. Both jumping to ν_1 (then choosing the safer action) and jumping to ν_2 (then choosing the riskier action) are optimal. When μ further increases, it becomes strictly optimal to choose a safer action, as depicted in Figure 8-(d). Figure 8-(e) shows that when μ is close to the upper bound of a continuation region, G becomes so concave that the optimal posterior ν is close to the prior, approximating a Gaussian signal. An extra "convex" region of G exists to the right of ν in Figure 8-(d)&(e). However, since the value that defines the multiplier in G is $V(\mu)$, strictly below F in the region, experimentation is dominated by stopping. In Figure 8-(f), for even higher μ and more concave G, the "convex" region remains; hence, it is again optimal to continue, and concavifying G pins down the optimal posterior. Moreover, the optimal policy is unique except for two (zero measure) cases: (i) $\mu = \mu^*$ and ν's on either side of μ can be optimal; (ii) choosing two different actions following the posteriors is optimal.

4.2.2 *Sketched proof of Theorem 2*

I prove Theorem 2 via construction and verification. I conjecture that the optimal policy for Equation (3) takes the form of Theorem 2: a confirmatory signal associated with an immediate action. I first construct $V(\mu)$ and $\nu(\mu)$ via three steps:

- *Step 1.* Determine μ^*. Since $\mu^* \in \arg\min V$, except for the special case where V is strictly monotonic, μ^* is essentially the unique belief at which $V'(\mu^*) = 0$, and searching for posteriors on either side of μ^* is optimal. The HJB equation implies an equation for μ^*:

$$\sup_{\nu\le\mu^*,I}\frac{F(\nu)-\frac{C(I)}{I}(H(\mu^*)-H(\nu)+H'(\mu^*)(\nu-\mu^*))}{1+\frac{\rho}{I}(H(\mu^*)-H(\nu)+H'(\mu^*)(\nu-\mu^*))}$$
$$=\sup_{\nu\ge\mu^*,I}\frac{F(\nu)-\frac{C(I)}{I}(H(\mu^*)-H(\nu)+H'(\mu^*)(\nu-\mu^*))}{1+\frac{\rho}{I}(H(\mu^*)-H(\nu)+H'(\mu^*)(\nu-\mu^*))}$$

$V(\mu^*)$ and $\nu(\mu^*)$ are pinned down correspondingly.

- *Step 2*. Solve for the value function while holding the action fixed. Let a be the optimal action for optimal posterior ν solved in step 1. Let $F_a(\mu)=E_\mu[u(a,x)]$. Now, solve for the value function given the payoff $F_a(\nu)$:

$$\rho V(\mu)=\max_{\nu\ge\mu,I}-I\frac{F_a(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{H(\nu)-H(\mu)-H'(\mu)(\nu-\mu)}-C(I)$$

  The primitives in the objective function are all sufficiently differentiable in ν. The first-order condition w.r.t. ν and I then yields a well-behaved ODE characterizing $\nu(\mu)$ and $I(\mu)$ with initial condition $\nu(\mu^*)$ and $I(\mu^*)$. Therefore, we can solve for the optimal policy (ν,I) and calculate value $V(\mu)$ accordingly for $\mu\ge\mu^*$. $V(\mu)$, $\nu(\mu)$ and $I(\mu)$ for all $\mu\le\mu^*$ are solved using a symmetric method.

- *Step 3*. Update the value function w.r.t. all alternative actions and smoothly paste the solved value function piece by piece. This step begins with solving the ODE defined in step 2 at μ^*. Next, I extend the value function toward $\mu=\{0,1\}$. Whenever a belief at which two actions yield the same payoff is reached, the value function is pasted to the solution of a new ODE corresponding to the new action. This process continues until the value function $V(\mu)$ smoothly pastes to $F(\mu)$. This procedure generates a quasi-convex value function (minimized at μ^*).

Solving the ODE characterizing $(\nu(\mu),I(\mu))$ directly implies the monotonicity of $\nu(\mu)$ and $I(\mu)$ in each connected experimentation region. At this point, the optimality of the constructed strategy requires verification, which takes three steps to rule out repeated jumps, contradictory evidence, and Gaussian signals. Section 4.2 explained the intuition for the suboptimality of the three alternative strategies. The formal proof is relegated to Appendix B.

5 Discussions and extensions

In this section, I discuss the assumptions. Section 5.1 illustrates the implications of the information measures defined by Assumption 1. Section 5.1 also presents predictions under alternative information measures, including "prior invariant" ones. In

Section 5.2, I relax Assumption 2 to allow for linear information cost functions. In Section 5.3, I relax Assumption 3 to allow for more than two states. Then, in Section 5.4 I discuss the role of exponential discounting.

5.1 *Information measure*

The assumption on the flow information measure $\langle I_t \rangle$ plays a crucial role in the model, as it dictates how strategies compare regarding their costs. In what follows, I first explain the implication of Assumption 1: it equalizes the average measures of different dynamic strategies that acquire the same information and incur the same expected delay; hence, it minimizes the impact of cost on the comparison between different strategies that acquire the same information. I then present extensions of the baseline model.

Uniform posterior separability

An equivalent statement of Assumption 1 is that the flow measure of information is a *uniformly posterior separable* (UPS) function of the conditional belief distribution, where the UPS property is defined below. Let X be a finite set, and let $\Delta^2(X)$ denote probability measures on $\Delta(X)$.

Definition 1. *$f : \Delta^2(X) \to \mathbb{R}^+$ is* ***uniformly posterior separable*** *if there exists a concave function $H : \Delta(X) \to \mathbb{R}$ s.t. $\forall \pi \in \Delta^2(X)$ and $\mu = \mathbb{E}_{\pi(\nu)}[\nu]$,*

$$f(\pi) = \mathbb{E}_{\pi(\nu)}[H(\mu) - H(\nu)].$$

Recall that Assumption 1 can be rewritten as $I_t = \lim_{s \to 0} \frac{1}{s} f(\mu_{t+s} | \mathcal{F}_t)$, for a UPS function $f(\cdot)$.[15] UPS functions and their special cases like mutual information and quadratic variation have been widely used to model the cost of information in models with flexible information acquisition. Their implications in static decision problems have been extensively studied (Caplin and Dean (2013), Caplin, Dean, and Leahy (2017), Frankel and Kamenica (2019), and Matějka and McKay (2014)). In what follows, I introduce a novel implication of UPS in dynamic settings. For a better exposition, the analysis is in a discrete-time setting; it incorporates "hat" to all discrete-time stochastic processes for differentiation. Let $\widehat{\mathbb{M}}$ denote the collection of discrete-time martingales on $\Delta(X)$.

Definition 2. *$f : \Delta^2(X) \to \mathbb{R}^+$ is* ***path independent*** *if there exists $g : \Delta^2(X) \times [1, \infty) \to \mathbb{R}^+$ s.t. $\forall \langle \widehat{\mu}_t \rangle \in \widehat{\mathbb{M}}$ and stopping time $\widehat{\tau}$,*

$$\frac{1}{\mathbb{E}[\widehat{\tau}]} \cdot \mathbb{E}\left[\sum_{t=1}^{\widehat{\tau}} f(\widehat{\mu}_t | \widehat{\mathcal{F}}_{t-1})\right] = g(\widehat{\mu}_{\widehat{\tau}}, \mathbb{E}[\widehat{\tau}]). \tag{9}$$

[15] To simplify notations, when a function takes a probability measure as its argument, I sometimes use random variables to denote their distributions. In $I(\mu_{t+s} | \mathcal{F}_t)$, $\mu_{t+s} | \mathcal{F}_t$ denotes the distribution of μ_{t+s} conditional on $\mathcal{F}_t$.

The LHS of Equation (9) is the average information measure of strategy $(\widehat{\mu}_t, \widehat{\tau})$. The RHS of Equation (9) is a function that depends on the distribution of random variable $\widehat{\mu}_{\widehat{\tau}}$—the overall information acquired via the strategy—and the expected delay $\mathbb{E}[\widehat{\tau}]$. Path independence says that the average information measure only depends on the overall information acquired and the expected delay, but not the intermediate steps specified by process $\langle \widehat{\mu}_t \rangle$. Theorem 3 shows that UPS is essentially equivalent to path independence. Let $\delta_\mu \in \Delta^2(X)$ denote point mass on μ.

Theorem 3. *$f : \Delta^2(X) \to \mathbb{R}^+$ is path independent and $\forall \mu \in \Delta(X)$, $f(\delta_\mu) = 0 \iff f$ is UPS.*

Proof. See Section S3.1 of the supplemental material. □

Theorem 3 shows that Assumption 1 guarantees that the flow measure of information is path independent.[16] The implication of path independence can be understood by thinking of the DM's decision problem as a hypothetical two-stage problem. In the first stage, the DM chooses the overall informativeness of the information to be acquired and the expected delay. In the second stage, the DM chooses how the information is acquired over time. The path independence property ensures that the exact specification of the information cost mainly affects the first-stage problem. In the second stage, the signal process does not affect the average flow measure of information. Therefore, the choice of the intertemporal distribution of information is mainly determined by how information benefits decision-making dynamically.[17] In particular, my model reveals that the optimal signal process is pinned down by studying the evolution of the precision-frequency tradeoff. This intuition is supported by the results of Hébert and Woodford (2018). They employ a similar model except for their assumption of constant delay cost instead of discounting; hence, the DM's expected payoff only depends on the overall information acquired and the expected delay. In this knife-edge case, the second stage problem is trivial: virtually all signal processes are payoff equivalent because of path independence. Thus, the dynamic aspects of the signal process are irrelevant, and the model reduces to a static rational inattention model.

A similar path independence property is observed by Morris and Strack (2019). Their result can be interpreted as follows: when restricting the domain to only one-dimensional Gaussian belief martingales, all flow information measures are path in-

[16] The equivalence between path independence and UPS (which claims the existence of a potential function H) is similar in form to the Poincaré lemma and hence may be expected.

[17] Formally, although the average flow information measure is fixed, the exact cost process may still be payoff relevant due to a nontrivial cost function $C(\cdot)$ and discounting. Therefore, while Assumption 1 largely limits the interference of cost on dynamic learning, it is not fully eliminated.

dependent. The key observation is that for any flow information measure defined on Gaussian processes, there exists a UPS extension for general signal processes. Path independence then follows.[18].

While we see that Assumption 1 is a benchmark that compares information "impartially" in dynamic environments, the cost of information can take other forms in practices, which may lead to different predictions. A leading critique of the UPS model is the "prior-invariance critique". It states that the cost of information should not depend on the subjective belief of a DM in many realistic settings where information takes the form of a physical device or physical evidence. Next, I extend my analysis to non-UPS information measures, including prior invariant ones.

Prior invariant information measure

To formally define the prior-invariance property, it is useful to model information via *statistical experiments*. A statistical experiment specifies a signal space S and state contingent signal distributions $p_x(s)$. An information measure is *prior invariant* if it is a function of the statistical experiment but not the prior belief. Gentzkow and Kamenica (2014) note that prior invariance is incompatible with UPS. Several recent papers have studied the cost of information in a prior invariant framework (Denti, Marinacci, Rustichini, et al. (2019), Mensch (2018), and Pomatto, Strack, and Tamuz (2020)). While the implication of a general prior invariant information measure in my model remains an open question, I illustrate through examples that replacing Assumption 1 with different prior invariant measures might strengthen or weaken the main results.[19]

Example 3. Let the information measure of the Poisson signal (p, ν) at belief μ be $p \cdot J(\mu, \nu) = p \cdot (\frac{\nu}{\mu} + \frac{1-\nu}{1-\mu}) g\left(\frac{\nu}{1-\nu}\frac{1-\mu}{\mu}\right)$, where $g \in C^2\mathbb{R}^+$. First, to see that this measure is prior invariant, suppose that a signal appears with Poisson rate λ_0 and λ_1 in the two states, respectively and induces posterior belief ν. Bayes rule implies that $p \cdot J(\mu, \nu) = p \cdot \left(\frac{\lambda_1}{p} + \frac{\lambda_0}{p}\right) g\left(\frac{\lambda_1}{p} \cdot \frac{p}{\lambda_0}\right) = (\lambda_0 + \lambda_1) g(\frac{\lambda_1}{\lambda_0})$, which is prior invariant.

Since Gaussian signals can be viewed as the uninformative limit of Poisson signals, the corresponding measure of Gaussian signals per unit variance is $\lim_{\nu \to \mu} \frac{J(\mu,\nu)}{(\nu-\mu)^2} = \frac{g''(1)}{(\mu-\mu^2)^2}$. Holding $g''(1)$ constant, the function $g(\cdot)$ can be chosen sufficiently steep elsewhere such that any sufficiently precise Poisson signal incurs a high cost that it cannot be optimal. Formally, Proposition 1 quantifies how steep $g(\cdot)$ need to be to rule out high precision Poisson signals:

[18] Morris and Strack (2019) employ a continuous-time setting, different from the discrete-time setting in Theorem 3. Thus, there is no formal inclusion relationship between the results.

[19] Formally defining prior invariant information measure on general belief martingales in continuous time is tricky. In the examples I consider only jump diffusion processes, on which prior invariance is well defined.

Proposition 1. *Given F and C satisfying Assumption 3, there exists $\Delta \in \mathbb{R}^+$ s.t. if $V \in C^1[0,1] \cap C^2(E)$ solves*

$$\rho V(\mu) = \max\left\{\rho F(\mu), \sup_{p,\nu,\sigma^2} p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) + \tfrac{1}{2}\sigma^2 V''(\mu)\right\}$$
$$- C\left(p \cdot \left(\tfrac{\nu}{\mu} + \tfrac{1-\nu}{1-\mu} g\left(\tfrac{\nu}{1-\nu}\tfrac{1-\mu}{\mu}\right)\right) + \tfrac{1}{2}\sigma^2 \tfrac{g''(1)}{(\mu-\mu^2)^2}\right),$$

then policy $(p > 0, \nu)$ is optimal at μ only if $g\left(\frac{\nu}{1-\nu}\frac{1-\mu}{\mu}\right) < \Delta g''(1)(H^*(\nu) - H^*(\mu) - H^{*\prime}(\mu)(\nu-\mu))$, where $H^*(\mu) = (2\mu - 1)\log\left(\frac{\mu}{1-\mu}\right)$.

Proof. See Section S3.2 of the supplemental material. □

Proposition 1 suggests one way to construct prior invariant information measures under which the optimal strategy is approximately Gaussian. Suppose for small $\epsilon > 0$, the function g is constructed to be steeper than the Bregman divergence measure defined by H^* multiplied a sufficiently large constant for any posterior belief outside the ϵ-neighborhood around the prior. Then the optimal Poisson signal must induce a posterior within the ϵ-neighborhood. Hence, the optimal belief process involves approximately continuous paths.

Example 4. Let the information measure of the Poisson signal (p, ν) at belief μ be $p \cdot J(\mu, \nu) = p \cdot \frac{|\nu-\mu|}{\mu(1-\mu)}$. Suppose that a signal appears with Poisson rate λ_0 and λ_1 in the two states, respectively and induces posterior belief ν. Bayes rule implies that $p \cdot J(\mu,\nu) = p \cdot |\frac{\nu}{\mu} - \frac{1-\nu}{1-\mu}| = p \cdot |\frac{\lambda_1}{p} - \frac{\lambda_0}{p}| = |\lambda_1 - \lambda_0|$, which is prior invariant.

Since $J(\mu,\nu)$ is in the order of $|\nu - \mu|$, the corresponding measure of Gaussian signals per unit variance $\lim_{\nu\to\mu} \frac{J(\mu,\nu)}{(\nu-\mu)^2} = \infty$. Thus, a Gaussian signal is never optimal when V'' exists since its benefit per unit variance $\frac{1}{2}V''(\mu)$ is finite. Moreover, this information measure also excludes low precision Poisson signals, as shown by Proposition 2.

Proposition 2. *Given F and C satisfying Assumption 3, suppose convex function $V \in C^1[0,1]$ solves*

$$\rho V(\mu) = \max\left\{\rho F(\mu), \sup_{p,\nu} p(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)) - C\left(p \cdot \tfrac{|\nu-\mu|}{\mu(1-\mu)}\right)\right\}.$$

Then policy $(p > 0, \nu)$ is optimal at μ only if $\nu \in E^C$.

Proof. See Section S3.3 of the supplemental material. □

Proposition 2 states that there exists prior invariant information measures under which the optimal Poisson signals are so precise that the induced posteriors are in the stopping region, which is exactly property (v) of Theorem 2. Examples 3 and 4 illustrates that the prior-invariance property is sufficiently general to accommodate both settings that strengthen the optimality of Poisson learning and settings that weaken it. The intuition is exactly from Theorem 3: in the two examples, I distort the information measures so that processes with large jumps incur higher and lower cost, respectively, relative to processes with small jumps. When such distortion is strong enough, it outweighs other incentives and dictates the optimal strategy.

General information measure

In this section, I extend the model to nearly fully general information measures. Consider the following extension of HJB Equation (3):

$$\rho V(\mu) = \max \left\{ \rho F(\mu), \sup_{p,\nu,\sigma^2} p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) + \frac{1}{2}\sigma^2 V''(\mu) \right\} \tag{10}$$
$$- C(pJ(\mu,\nu) + \kappa(\mu,\sigma)),$$

where $pJ(\mu,\nu)$ and $\kappa(\mu,\sigma)$ define the flow information measure of the Poisson signal and Gaussian signal, respectively. To obtain tractability under such generality, I directly work with the HJB equation and consider one Poisson signal and one Gaussian signal in the optimization problem. The information measure satisfies the following regularity assumption.

Assumption 4.

(i) $\forall \mu \in (0,1)$, $J(\mu,\cdot) \in C^2(0,1)$, $J(\mu,\mu) = J'_\nu(\mu,\mu) = 0$, *and* $J''_{\nu\nu}(\mu,\mu) > 0$.

(ii) $\kappa(\mu,\sigma) = \frac{1}{2}\sigma^2 J''_{\nu\nu}(\mu,\mu)$.

Part (i) of Assumption 4 states that "any uninformative signal is free" and "any informative signal is costly". Part (ii) states that the information measure is "continuous" when a Poisson signal approximates a Gaussian signal. Define increasing function $h(x)$ implicitly through $h(x)C'^{-1}(h(x)) - C(C'^{-1}h(x)) = x$.

Theorem 4. *Given F and C satisfying Assumption 3 and* $J \in C^4[0,1]^2$ *satisfying Assumption 4, suppose* $V \in C^1[0,1] \cap C^3(E)$ *solves Equation (10). Let* $L(\mu,x)$ *be defined by*

$$L(\mu,x) = \rho h'(x) J''_{\nu\nu}(\mu,\mu)^2 - \frac{h''(x)h(x)}{h'(x)^2} \frac{J^{(3)}_{\nu\nu\mu}(\mu,\mu)^2}{J''_{\nu\nu}(\mu,\mu)^2}$$
$$- \frac{2J^{(3)}_{\nu\nu\mu}(\mu,\mu)^2 + J^{(3)}_{\nu\nu\mu}(\mu,\mu)J^{(3)}_{\nu\nu\nu}(\mu,\mu)}{J''_{\nu\nu}(\mu,\mu)} + J^{(4)}_{\nu\nu\mu\mu}(\mu,\mu) + J^{(4)}_{\nu\nu\nu\mu}(\mu,\mu).$$

Then in the open region $D = \left\{\mu \in E \middle| \forall x \in [\rho \min F, \rho \max F], L(\mu, x) \neq 0\right\}$*, the set of* μ *s.t.*

$$\rho V(\mu) = \max_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - C(\kappa(\mu, \sigma))$$

is of zero measure.

Proof. See Section S3.4 of the supplemental material. □

Theorem 4 states that when the information measure J satisfies a collection of differential inequalities ($L(\mu, x) \neq 0$), Gaussian learning is strictly dominated almost everywhere. Loosely speaking, the region D is often a non-degenerate set, because the set of functions that satisfy a differential equation is typically non-generic. The suboptimality of Gaussian learning is primarily because the continuity condition in Assumption 4 guarantees that the path-independence property is satisfied locally for low precision Poisson and Gaussian signals, although high precision Poisson signals can be arbitrarily costly. Thus, the key precision-frequency tradeoff remains the same locally; hence, Gaussian signals are dominated by low precision Poisson signals.[20]

The differential inequality condition is quite involved; however, it allows for generating simple sufficient conditions for the suboptimality of Gaussian learning.

Corollary 1. Fixing functions F, J and $C(I) = I^{\alpha}$, let the conditions in Theorem 4 be satisfied. $\exists \underline{\rho}$ s.t. if $\rho \geq \underline{\rho}$ and V solves Equation (10), then the set of $\mu \in E$ s.t. $\rho V(\mu) = \max_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - C(\kappa(\mu, \sigma))$ is of zero measure.

Proof. See Section S3.4 of the supplemental material. □

The sufficient condition in corollary 1 states that the information cost C is homogeneous and the DM is sufficiently impatient.

Corollary 2. Fixing functions F, κ and C, let the conditions in Theorem 4 be satisfied. $\exists \varepsilon > 0$ s.t. if $\left\| J^{(3)}_{\nu\nu\mu}(\mu, \mu) \right\| + \left\| \nabla J^{(3)}_{\nu\nu\mu}(\mu, \mu) \right\| \leq \varepsilon$ and V solves Equation (10), then the set of $\mu \in E$ s.t. $\rho V(\mu) = \max_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - C(\kappa(\mu, \sigma))$ is of zero measure.

Proof. See Section S3.4 of the supplemental material. □

The sufficient condition in corollary 2 states that the information measure is perturbed near Assumption 1, which requires $J^{(3)}_{\nu\nu\mu}(\mu, \mu) \equiv 0$. The results of corollaries 1 and 2 are intuitive: when the incentives provided by the dynamic nature of decision-making outweigh the distortion provided by the information measure, the same force as in the main model renders Gaussian learning suboptimal.

[20] Note that Theorem 4 is consistent with with Proposition 1 because the size of the jumps can be arbitrarily small.

5.2 Linear cost function

The convexity condition in Assumption 2 is a standard increasing-marginal-cost assumption. Its main implication is that the DM has an incentive to smooth the cost over time. When the cost of information $C(I)$ is a linear function, the optimal value is achieved by immediately acquiring all the information and making a decision, as per Theorem 5.

Theorem 5. *Given Assumption 1 and $C(I) = \lambda I$, suppose $V(\mu)$ solves Equation (1); then,*

$$V(\mu) = \sup_{\mathbb{E}_\pi[\nu]=\mu} E_\pi[F(\nu)] - \lambda E_\pi[H(\mu) - H(\nu)] \tag{11}$$

Proof. See Section S3.5 of the supplemental material. □

The RHS of Equation (11) is the payoff from acquiring π at cost $\lambda(H(\mu) - \mathbb{E}_\pi[H(\nu)])$ and deciding immediately. Theorem 5 states that this strategy is optimal. The intuition is simple. Since the marginal cost is constant, consolidating all information acquired in the future to the first period does not affect the cost incurred but strictly reduces the utility loss from discounting.

In fact, given Assumption 1 and linear $C(I)$, Equation (1) is a variant of the general model studied by Steiner, Stewart, and Matějka (2017), which further considers a varying state and repeated decision making. This analysis clarifies that the dynamics in Steiner, Stewart, and Matějka (2017) are a result of the intertemporal dependence of decision problems instead of strategically delaying the decision to acquire information.

5.3 General state space

The binary state assumption in Assumption 3 is technically crucial for my analysis. Because it reduces the belief space $\Delta(X)$ to a one-dimensional interval; hence, ODE theory can be applied to obtain the solution to the HJB equation.[21] However, the key tradeoffs and intuitions I develop in Section 4.2 do not rely on the binary state assumption. This claim is verified in Theorem 6, which shows that with more than two states, if a solution to the HJB equation exists, the qualitative properties of the optimal strategy all generalize.

Theorem 6. *Given C satisfying Assumption 3 and strictly concave $H \in C^2\Delta(X)$. Suppose quasi-convex $V \in C^1\Delta(X) \cap C^2(E)$ solves Equation (3). Then, $\exists$ policy function $(\nu(\mu), p(\mu))$ on E satisfying*

$$\rho V(\mu) = p(\mu)(F(\nu(\mu)) - V(\mu) - \nabla V(\mu)(\nu(\mu) - \mu))$$

[21] The main challenge of extending the analysis to multidimensional belief spaces using PDE theory is that the boundary conditions are endogenous. They can be easily guessed in the one-dimensional case (step 1 of Section 4.2.2) but are much harder to guess generally.

$$- C(-p(\mu)(H(\nu(\mu)) - H(\mu) - \nabla H(\mu)(\nu(\mu) - \mu)))$$

and the following properties:

(i) *Poisson learning:* $\rho V(\mu) \geq \max_{\sigma} \frac{1}{2}\sigma^T \text{Hess} V(\mu)\sigma - C\left(-\frac{1}{2}\sigma^T \text{Hess} H(\mu)\sigma\right)$.

(ii) *Direction:* $D_{\nu(\mu)-\mu} V(\mu) \geq 0$.

(iii) *Precision:* $(\nu(\mu) - \mu)^T(-\text{Hess} H(\nu)) D_{\nu(\mu)-\mu}\nu(\mu) \leq 0$ *when* ν *is differentiable at* μ.

(iv) *Intensity:* $I(\mu) \triangleq p(\mu)(H(\mu) - H(\nu(\mu)) + \nabla H(\mu)(\nu(\mu) - \mu))$ *is increasing in* $V(\mu)$.

(v) *Stopping time:* $\nu(\mu) \in E^C$.

There exists a nowhere dense set K *s.t. strict inequality holds on* $E \setminus K$ *for properties 1,2, and 3.*

Proof. See Section S3.6 of the supplemental material. □

Theorem 6 is an analogy of Theorem 2 under the stronger assumption that a value function solving Equation (3) exists. The value function is implemented by a Poisson signal satisfying five properties. Property (i) states that Gaussian signals are strictly suboptimal except for a nowhere dense set of beliefs. Property (ii) states that it is optimal to jump toward the direction where the value function increases. Property (iii) states that when prior belief moves toward the optimal posterior, the optimal posterior moves toward the prior, under proper rotation and scaling of the linear space $\Delta(X)$.[22] Property (iv) states that the optimal intensity is positively related to the continuation value. Property (v) states that it is optimal to immediately stop upon the signal's arrival. Evidently, when $|X| = 2$, the five properties precisely reduce to the corresponding ones in Theorem 2. When $|X| > 2$, properties (i),(iv) and (v) almost hold identically, while properties (ii) and (iii) are high-dimension extensions of the corresponding properties in Theorem 2.

Therefore, the optimal learning dynamics is qualitatively the same as in the binary state case. The DM seeks a Poisson signal whose arrival "confirms" the prior belief by leading the posterior to jump toward the direction where the value function increases. The signal is sufficiently precise to warrant immediate action. Absent the signal, the DM's value function decreases over time, leading to lower spending in learning. In a special case when $H(\mu)$ has a constant Hessian matrix, the Bregman divergence between μ and ν increases over time, indicating increasing precision and decreasing frequency.[23]

[22] Let $U^T \Lambda U$ be the spectral decomposition of $\text{Hess} H(\nu)$. Property (iii) of Theorem 6 means that $\nu(\mu) - \mu$ and $\nabla_{\nu(\mu)-\mu}\nu(\mu)$ have a negative inner product when both are rotated by U and scaled by $\sqrt{-\Lambda}$.

[23] Let $J(\mu) = H(\mu) - H(\nu(\mu)) + \nabla H(\mu)(\nu(\mu) - \mu)$ (the Bregman divergence associated with $-H$). In this case, $D_{\mu-\nu(\mu)} J(\mu) = -(\nu(\mu) - \mu)^T \cdot \text{Hess} H \cdot D_{\mu-\nu(\mu)}\nu(\mu) - (\nu(\mu) - \mu)^T \cdot \text{Hess} H \cdot (\nu(\mu) - \mu)$. Property (iii) of Theorem 6 implies $D_{\mu-\nu(\mu)} J(\mu) \geq 0$. Decreasing frequency follows from decreasing $I(\mu)$.

5.4 *Exponential discounting*

My model employs the expected discounted utility (EDU) framework to model the DM's time preference, arguably the standard approach in economics. Interestingly, the form of time preference has sharp implications for information acquisition. The strategy of the DM pins down a joint distribution of stopping time and payoffs. How the DM evaluates the distribution in the dimension of stopping time depends crucially on her time preference. In the standard EDU framework, the DM assigns utility $e^{-\rho t}$ to each unit of period t payoffs. Since the exponential function is strictly convex, an EDU maximizer is effectively "risk loving" in the time dimension.[24] Meanwhile, the Poisson learning strategy introduced in Theorem 2 involves high "time risk". Since the strategy only seeks signals precise enough to lead to stopping, the strategy effectively maximizes the probability of an early decision at the cost of possible long delays (learning becomes increasingly difficult over time because belief drifts away from the decision boundaries). The resulting decision time is highly dispersed. The Gaussian learning strategy, on the other hand, involves lower time risk, because the typical hitting time of a diffusion process at absorbing boundaries is in a less dispersed hump shape.

In certain settings, the DM may exhibit alternative time preferences. For example, a fixed flow waiting cost specifies a time-risk neutral preference. An exogenous deadline specifies a time-risk averse preference. As discussed in Section 5.1, Hébert and Woodford (2018)'s model studies the time-risk neutral setting, in which the path independence property makes virtually all types of strategies, including Gaussian learning, payoff equivalent. Regarding the time-risk averse case, the "suspense maximizing" strategy introduced by Ely, Frankel, and Kamenica (2015) is a candidate solution since it leads to a deterministic stopping time.[25] It is worth noting that since Gaussian learning does not minimize the time risk either, it is generally suboptimal except in knife-edge cases.

6 Conclusion

This study provides a dynamic information acquisition framework that allows for a fully flexible design of the signal processes. Under the assumptions that the DM

[24] DeJarnette et al. (2020) formally introduce a notion of time-risk attitudes and point out the implications of EDU for time-risk attitudes. They provide lab evidence against time-risk loving and generalizes the EDU framework to accommodate non-time-risk loving preferences.

[25] Ely, Frankel, and Kamenica (2015) construct a discrete-time belief martingale with constant quadratic variation rate that fully reveals the state at a deterministic time. Lemma S.1 in Section S1 of the supplemental material extends their construction to a continuous-time belief process with constant uncertainty reduction rate that attains an arbitrary distribution of belief at a deterministic time.

discounts future utility and the information cost depends on the uncertainty reduction speed, I fully characterize the optimal policy: the optimal information structure induces beliefs following a compensated Poisson process. The arrival of the Poisson signal confirms the most promising action at the moment and leads to an immediate action. The absence of the signal is followed by continued learning with increasing precision and decreasing frequency.

Using the main model as a benchmark, the discussions on the key assumptions raise two questions for future research. First, I illustrate that if the cost of information is "prior invariant", the optimality of Poisson signals may or may not hold, depending on the specification of the cost. The general implication of prior invariant information cost for dynamic information acquisition remains an open question. Second, my model reveals that the seemingly standard expected discounted utility framework has strong implications for dynamic information acquisition via its induced risk attitude toward time. It would be interesting to systematically understand the "time-risk" involved in the dynamic learning strategies.

References

Arrow, Kenneth J, David Blackwell, and Meyer A Girshick (1949). "Bayes and minimax solutions of sequential decision problems." In: *Econometrica, Journal of the Econometric Society*, pp. 213–244.

Aumann, Robert J, Michael Maschler, and Richard E Stearns (1995). *Repeated games with incomplete information*. MIT press.

Bergemann, Dirk and Juuso Välimäki (1996). "Learning and strategic pricing." In: *Econometrica: Journal of the Econometric Society*, pp. 1125–1149.

Boel, René and Michael Kohlmann (1980). "Semimartingale models of stochastic optimal control, with applications to double martingales." In: *SIAM Journal on Control and Optimization* 18.5, pp. 511–533.

Bolton, Patrick and Christopher Harris (1999). "Strategic experimentation." In: *Econometrica* 67.2, pp. 349–374.

Callander, Steven (2011). "Searching and learning by trial and error." In: *American Economic Review* 101.6, pp. 2277–2308.

Caplin, Andrew and Mark Dean (Aug. 2013). *Behavioral Implications of Rational Inattention with Shannon Entropy*. Working Paper 19318. National Bureau of Economic Research. DOI: 10.3386/w19318. URL: http://www.nber.org/papers/w19318.

Caplin, Andrew, Mark Dean, and John Leahy (Aug. 2017). *Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy*. Working Paper 23652. National Bureau of Economic Research. DOI: 10.3386/w23652. URL: http://www.nber.org/papers/w23652.

Che, Yeon-Koo and Konrad Mierendorff (2019). “Optimal Dynamic Allocation of Attention.” In: *American Economic Review* 109.8, pp. 2993–3029.

Clark, Aubrey and Giovanni Reggiani (2021). *Contracts for acquiring information*. arXiv: 2103.03911 [econ.TH].

Davis, Mark HA (1979). “Martingale methods in stochastic control.” In: *Stochastic Control Theory and Stochastic Differential Systems*. Springer, pp. 85–117.

DeJarnette, Patrick et al. (2020). “Time lotteries and stochastic impatience.” In: *Econometrica* 88.2, pp. 619–656.

Denti, Tommaso, Massimo Marinacci, Aldo Rustichini, et al. (2019). *Experimental cost of information*. Tech. rep.

Doval, Laura (2018). “Whether or not to open Pandora’s box.” In: *Journal of Economic Theory* 175, pp. 127–158.

Ely, Jeffrey (2017). “Beeps.” In: *American Economic Review* 107.1, pp. 31–53.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2015). “Suspense and surprise.” In: *Journal of Political Economy* 123.1, pp. 215–260.

Frankel, Alexander and Emir Kamenica (2019). “Quantifying information and uncertainty.” In: *American Economic Review* 109.10, pp. 3650–80.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018). “Speed, accuracy, and the optimal timing of choices.” In: *American Economic Review* 108.12, pp. 3651–84.

Gentzkow, Matthew and Emir Kamenica (2014). “Costly persuasion.” In: *American Economic Review* 104.5, pp. 457–62.

Gittins, John (1974). “A dynamic allocation index for the sequential design of experiments.” In: *Progress in statistics*, pp. 241–266.

Hébert, Benjamin and Michael Woodford (Nov. 2018). *Information Costs and Sequential Information Sampling*. Working Paper 25316. National Bureau of Economic Research. DOI: 10.3386/w25316. URL: http://www.nber.org/papers/w25316.

Kamenica, Emir and Matthew Gentzkow (2011). “Bayesian persuasion.” In: *American Economic Review* 101.6, pp. 2590–2615.

Ke, T Tony and J Miguel Villas-Boas (2019). “Optimal learning before choice.” In: *Journal of Economic Theory* 180, pp. 383–437.

Klabjan, Diego, Wojciech Olszewski, and Asher Wolinsky (2014). “Attributes.” In: *Games and Economic Behavior* 88, pp. 190–206.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis (2018). “Optimal and Myopic Information Acquisition.” arXiv: 1703.06367 [cs.GT].

Matějka, Filip and Alisdair McKay (2014). “Rational inattention to discrete choices: A new foundation for the multinomial logit model.” In: *The American Economic Review* 105.1, pp. 272–298.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva (2019 forthcoming). "On information design in games." In: *Journal of Political Economy*.
Matyskova, Ludmila (2018). "Bayesian persuasion with costly information acquisition." In: *CERGE-EI Working Paper Series* 614. URL: https://dx.doi.org/10.2139/ssrn.3161174.
Mayskaya, Tatiana (2020). "Dynamic Choice of Information Sources." URL: https://dx.doi.org/10.2139/ssrn.2863605.
Mensch, Jeffrey (2018). "Cardinal representations of information." In: URL: https://dx.doi.org/10.2139/ssrn.3148954.
Morris, Stephen and Philipp Strack (2019). *The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs*. Working Paper. URL: https://ssrn.com/abstract=2991567.
Moscarini, Giuseppe and Lones Smith (2001). "The optimal level of experimentation." In: *Econometrica* 69.6, pp. 1629–1644.
Pomatto, Luciano, Philipp Strack, and Omer Tamuz (2020). *The cost of information*. arXiv: 1812.04211 [econ.TH].
Rappoport, Daniel and Valentin Somma (2017). *Incentivizing Information Design*. URL: https://dx.doi.org/10.2139/ssrn.3001416.
Shannon, Claude Elwood (1948). "A mathematical theory of communication." In: *Bell system technical journal* 27.3, pp. 379–423.
Sims, Christopher A (2003). "Implications of rational inattention." In: *Journal of monetary Economics* 50.3, pp. 665–690.
Steiner, Jakub, Colin Stewart, and Filip Matějka (2017). "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." In: *Econometrica* 85.2, pp. 521–553.
Striebel, Charlotte (1984). "Martingale conditions for the optimal control of continuous time stochastic systems." In: *Stochastic processes and their applications* 18.2, pp. 329–347.
Wald, Abraham (1947). "Foundations of a general theory of sequential decision functions." In: *Econometrica, Journal of the Econometric Society*, pp. 279–313.
Weber, Richard et al. (1992). "On the Gittins index for multiarmed bandits." In: *The Annals of Applied Probability* 2.4, pp. 1024–1033.
Weitzman, Martin L (1979). "Optimal search for the best alternative." In: *Econometrica: Journal of the Econometric Society*, pp. 641–654.

Appendix: omitted proofs

A Proof of Theorem 1

To prove Theorem 1, I consider a sequence of auxiliary discrete-time problems that are discretizations of Equation (1). In each discrete-time problem, the primitives (A,X,u,μ,ρ) are the same as those in Section 3. Time is discrete $t\in\mathbb{N}$, and the period length $dt>0$. The payoff delayed by t periods is discounted by $e^{-\rho dt\cdot t}$. The DM chooses a belief martingale $\langle\widehat{\mu}_t\rangle$ and stopping time $\widehat{\tau}$. Let $\langle\widehat{\mathcal{F}}_t\rangle$ be the natural filtration of $\langle\widehat{\mu}_t\rangle$. The flow cost of information is $C\left(\frac{1}{dt}\mathbb{E}[H(\widehat{\mu}_t)-H(\widehat{\mu}_{t+1})|\widehat{\mathcal{F}}_t]\right)\cdot dt$. The DM solves the following stochastic control problem:

$$V_{dt}(\mu)=\sup_{\langle\widehat{\mu}_t\rangle\in\widehat{\mathbb{M}},\widehat{\tau}}\mathbb{E}\left[e^{-\rho dt\cdot\widehat{\tau}}F(\widehat{\mu}_{\widehat{\tau}})-\sum_{t=0}^{\widehat{\tau}-1}e^{-\rho dt\cdot t}C\left(\frac{1}{dt}\mathbb{E}\left[H(\widehat{\mu}_t)-H(\widehat{\mu}_{t+1})|\widehat{\mathcal{F}}_t\right]\right)\cdot dt\right] \quad (12)$$

where $\widehat{\mathbb{M}}$ is the set of discrete-time martingales satisfying $\widehat{\mu}_0=\mu$, and $\widehat{\tau}$ is a $\langle\widehat{\mathcal{F}}_t\rangle-$measurable stopping time. All discrete-time stochastic processes and random variables are labeled with "hat" to distinguish them from continuous-time processes. Lemma A.1 proves that the solutions of the auxiliary problems $V_{dt}(\mu)$ approximate that of the continuous-time problem when $dt\to0$.

Lemma A.1. *Given Assumptions 1 and 2,* $V_{dt}\xrightarrow[dt\to0]{l_\infty}V$.

Next, I characterize the solution of the auxiliary problem using dynamic programming theory.

Lemma A.2. *Given Assumptions 1 and 2, V_{dt} is the unique solution in $C(\Delta X)$ of the following functional equation:*

$$V_{dt}(\mu)=\max\left\{F(\mu),\max_{p,(\nu_i)}e^{-\rho dt}\sum_{i=1}^{2|X|}p_iV_{dt}(\nu_i)-C\left(\frac{H(\mu)-\sum p_iH(\nu_i)}{dt}\right)\cdot dt\right\} \quad (13)$$
$$s.t.\sum p_i\nu_i=\mu\ ,$$

where $p\in\Delta(N),\nu_i\in\Delta(X)$, $\forall i$.

Equation (13) is a standard Bellman equation, except for the choice of signal structure being restricted to have a support size no larger than $2|X|$, while Equation (12) permits an arbitrary number of signals. Lemma A.3 proves that whenever Equation (3) has a solution, the solution is unique and coincides with the limit of a sequence of solutions of the discrete-time problem Equation (13) when $dt\to0$.

Lemma A.3. *Given Assumptions 1 and 2, assume H is strictly concave and C^2 on interior beliefs. Suppose $V(\mu)\in C^1(\Delta(X))$ and $V_{dt}(\mu)\in C(\Delta(X))$ are solutions of Equations (3) and (13), respectively. Then,* $V_{dt}\xrightarrow[dt\to0]{l_\infty}V$.

Theorem 1 is a direct corollary of Lemmas A.1, A.2 and A.3, as is illustrated by Figure 9. The proofs of the three lemmas are provided in Appendices A.1, A.2 and A.3.

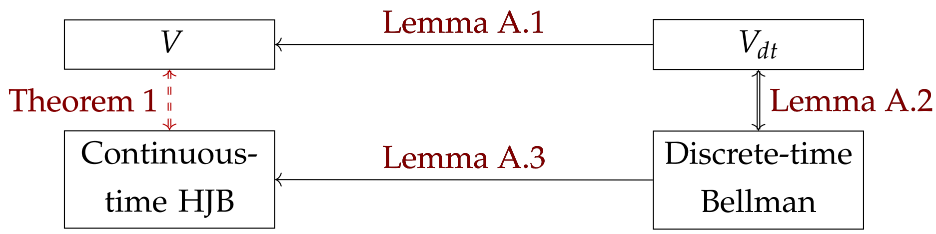

Figure 9: Proof of Theorem 1

A.1 *Proof of Lemma A.1*

Proof. First, the formal definition of the objective function in Equation (1) is as follows. For any admissible strategy $(\langle\mu_t\rangle,\tau)$, define a Riemann sum:

$$W_{dt}(\mu_t,\tau)=\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt])E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C(-\mathcal{A}_{jdt}H)dt\right].$$

Then, define the integral as

$$\mathbb{E}\left[e^{-\rho\tau}F(\mu_\tau)-\int_0^\tau e^{-\rho t}C(-\mathcal{A}_tH(\mu_t))\mathrm{d}t\right]\triangleq\varlimsup_{dt\to 0}W_{dt}(\mu_t,\tau). \tag{14}$$

Remark A.1. In Equation (14), the integral is defined as the limit superior. Clearly, defined this way, $V(\mu)$ may be strictly larger than what can be achieved when restricting to strategies such that the $\lim_{dt\to 0}W_{dt}(\mu_t,\tau)$ exists.[26] However, step 2 of the proof shows by construction that $V(\mu)$ can be achieved via strategies such that the limit exists. Therefore, imposing no further restriction on $(\langle\mu_t\rangle,\tau)$ is without loss. $\lim_{dt\to 0}V_{dt}$, on the other hand, always exists by the monotone convergence theorem.[27]

Step 1. Prove $V(\mu)\le\lim V_{dt}(\mu)$: Since $\langle\mu_t\rangle\in\mathscr{D}(H)$, $\langle\mathcal{A}_tH\rangle$ is a bounded process. Let the bound be M. Since $\langle H(\mu_t)\rangle$ is a semimartingale, $\forall s,t$, $-\mathcal{A}_t^sH=\mathbb{E}\left[\frac{1}{s}\int_t^{t+s}-\mathrm{d}H(\mu_t)\middle|\mathcal{F}_t\right]\le M$. Since $\langle\mathcal{A}_t^sH\rangle\xrightarrow{P}\langle\mathcal{A}_tH\rangle$, $\forall\varepsilon$ there exists $\delta>0$ s.t. $\forall s\le\delta$, $\text{Prob}(|\mathcal{A}_t^sH-\mathcal{A}_tH|_{l^\infty}\ge\varepsilon)\le\varepsilon$. Take any $dt\le\delta$, and let Ω denote the event $|\mathcal{A}_t^{dt}H-\mathcal{A}_tH|_{l^\infty}<\varepsilon$:

$$\begin{aligned}W_{dt}(\mu_t,\tau)=&\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt])E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C(-\mathcal{A}_{jdt}H)dt\right]\\&=\text{Prob}(\Omega)\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt]|\Omega)E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C(-\mathcal{A}_{jdt}H)dt\middle|\Omega\right]\end{aligned}$$

[26]When $\lim_{dt\to 0}W_{dt}(\mu_t,\tau)$ exists, Equation (14) defines the usual Riemann integral.

[27]A smaller dt gives the DM more flexibility in choosing the strategy. Thus, V_{dt} is an increasing sequence when $dt\to 0$. The detailed proof of this argument is relegated to Lemma S.3 in Section S1 of the supplemental material.

$$
\begin{aligned}
&+\text{Prob}(\Omega^C)\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt]|\Omega^C)E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C(-\mathcal{A}_{jdt}H)dt\middle|\Omega^C\right]\\
\leq&\text{Prob}(\Omega)\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt]|\Omega)E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C\left(-\mathcal{A}_{jdt}^{dt}H\right)dt\middle|\Omega\right]\\
&+\frac{C(M)-C(M-\varepsilon)}{1-e^{-\rho dt}}\\
&+\text{Prob}(\Omega^C)\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt]|\Omega^C)E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C\left(-\mathcal{A}_{jdt}^{dt}H\right)dt\middle|\Omega^C\right]\\
&+\varepsilon\cdot\frac{C(M)}{1-e^{-\rho dt}}\\
=&\sum_{i=1}^{\infty}\text{Prob}(\tau\in[(i-1)dt,idt])E\left[e^{-i\rho dt}F(\mu_{idt})-\sum_{j=0}^{i-1}e^{-j\rho dt}C\left(-\mathcal{A}_{jdt}^{dt}H\right)dt\right] \qquad (15)\\
&+\frac{C(M)-C(M-\varepsilon)}{1-e^{-\rho dt}}+\varepsilon\cdot\frac{C(M)}{1-e^{-\rho dt}}.
\end{aligned}
$$

The first equality is by definition. The second and third equalities are the law of iterated expectation. The inequality is implied by $C(\cdot)$ being convex and both $\mathcal{A}_t^s H$ and $\mathcal{A}_t H$ being bounded by M. Note that if $\langle\widehat{\mu}_t\rangle$ and $\widehat{\tau}$ are the discretizations of $\langle\mu_t\rangle,\tau$, then the RHS of Equation (15) is exactly the same as the objective function in Equation (2) under strategy $(\langle\widehat{\mu}_t\rangle,\widehat{\tau})$. Therefore, Equation (15) implies $\forall\varepsilon>0$, $\exists\delta>0$ s.t. $\forall dt\leq\delta$:

$$
\begin{aligned}
&W_{dt}(\mu_t,\tau)\leq V_{dt}(\mu)+\frac{C(M)-C(M-\varepsilon)}{1-e^{-\rho dt}}+\varepsilon\cdot\frac{C(M)}{1-e^{-\rho dt}}\\
\Longrightarrow&\overline{\lim_{dt\to 0}}W_{dt}(\mu_t,\tau)\leq\overline{\lim_{dt\to 0}}V_{dt}(\mu)\\
\Longrightarrow&V(\mu)=\sup_{\langle\mu_t\rangle,\tau}\overline{\lim_{dt\to 0}}W_{dt}(\mu_t,\tau)\leq\lim_{dt\to 0}V_{dt}(\mu).
\end{aligned}
$$

Step 2. Prove $V(\mu)\geq\lim V_{dt}(\mu)$: I prove this claim by showing that $\forall dt>0$, there exists a continuous-time strategy that achieves a payoff no less than $V_{dt}(\mu)$. Given period length dt and a discrete-time strategy $(\langle\widehat{\mu}_t\rangle,\widehat{\tau})$. The goal is to construct an admissible continuous-time belief process $\langle\mu_t\rangle$, which satisfies two properties: 1) the marginal distribution of $(\mu_{idt})_{i\in\mathbb{N}}$ is the same as that of $(\widehat{\mu}_i)_{I\in\mathbb{N}}$, and 2) $\forall i$, the uncertainty reduction speed of μ_t within $[idt,(i+1)dt]$ is $\mathbb{E}[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i]/dt$. Such $\langle\mu_t\rangle$ can be constructed as per Lemma S.1 (Section S1 of the supplemental material): $\forall i$ and conditional on $\widehat{\mathcal{F}}_i$, there exists a continuous-time martingale $\langle\widetilde{\mu}_t\rangle$ satisfying $\forall s,t\in[0,1]$, $s>t$: $\mathbb{E}[H(\widetilde{\mu}_t)-H(\widetilde{\mu}_s)|\widetilde{\mathcal{F}}_t]=(s-t)\mathbb{E}[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i]$. [28] For $t\in[idt,(i+1)dt]$, define

[28] Lemma S.1 is a technical extension of the discrete-time "suspense optimal" strategy introduced by Ely, Frankel, and Kamenica (2015) to continuous-time based on the Kolmogorov's theorem. The proof is relegated to the supplemental materials.

$\mu_t|\mathcal{F}_{idt}=\widetilde{\mu}_{\frac{t-idt}{dt}}|\widehat{\mathcal{F}}_i$. Therefore, $\forall t\in[idt,(i+1)dt)$:

$$-\mathcal{A}_t H(\mu_t)=\lim_{s\to t^+}\mathbb{E}\left[\frac{H(\mu_t)-H(\mu_s)}{s-t}\middle|\mathcal{F}_t\right]$$
$$=\lim_{s\to t^+}\frac{\frac{s-t}{dt}\mathbb{E}\left[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i\right]}{s-t}$$
$$=\frac{\mathbb{E}\left[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i\right]}{dt}.$$

Let $\tau=\widehat{\tau}dt$. By construction, τ is measurable to the natural filtration of $\langle\mu_t\rangle$. Evidently, $(\langle\mu_t\rangle,\tau)$ leads to the same decision payoff as $(\langle\widehat{\mu}_t\rangle,\widehat{\tau})$. Therefore,

$$V(\mu)\geq\mathbb{E}\left[e^{-\rho\tau}F(\mu_\tau)-\int_0^\tau e^{-\rho t}C(I_t)\mathrm{d}t\right]$$
$$=\mathbb{E}\left[e^{-\rho dt\cdot\widehat{\tau}}F(\widehat{\mu}_{\widehat{\tau}})-\sum_{t=0}^{\widehat{\tau}-1}C\left(\frac{\mathbb{E}\left[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i\right]}{dt}\right)e^{-\rho dt\cdot t}\cdot\frac{1-e^{-\rho dt}}{\rho}\right]$$
$$\geq\mathbb{E}\left[e^{-\rho dt\cdot\widehat{\tau}}F(\widehat{\mu}_{\widehat{\tau}})-\sum_{t=0}^{\widehat{\tau}-1}C\left(\frac{\mathbb{E}\left[H(\widehat{\mu}_i)-H(\widehat{\mu}_{i+1})|\widehat{\mathcal{F}}_i\right]}{dt}\right)e^{-\rho dt\cdot t}\cdot dt\right].$$

The second inequality is from $1-e^{-x}\leq x$. Since the inequality holds for all dt and strategy, $V(\mu)\geq\lim V_{dt}(\mu)$. □

A.2 Proof of Lemma A.2

Proof. *Step 1.* Verify the standard transversality condition. This step is trivial as the payoff is bounded by $\sup F$ and discounted exponentially.

Step 2. Verify the Blackwell contraction-mapping condition. The contraction parameter in Equation (13) is the discount factor $e^{-\rho dt}<1$. Let $\Delta^2(X)$ be equipped with the weak topology. $\forall dt$, $e^{-\rho dt}\mathbb{E}_\pi[V_{dt}(\nu)]-C((H(\mu)-\mathbb{E}_\pi[H(\nu)])/dt)dt$ is continuous in $\pi\in\Delta^2(X)$ because $\Delta(X)$ is compact and $V,H\in C(\Delta X)$. $\{\pi\in\Delta^2(X)|\mathbb{E}_\pi[\nu]=\mu\}$ is compact valued and continuous in μ by the Prokhorov's theorem. Then, Berge's maximum theorem applies; hence, the Blackwell mapping is a contraction and maps into $C(\Delta X)$.

Step 3. With steps 1-2, the contraction-mapping fixed point theorem implies a unique value function V_{dt} solving Equation (13) (without the restriction on the support size). Lemma S.4 from the supplemental material shows that the optimal strategy of Equation (13) can be restricted to a support size $2|X|$. [29] □

[29] Lemma S.4 extends a familiar result in Bayesian persuasion (Aumann, Maschler, and Stearns (1995) and Kamenica and Gentzkow (2011)).

A.3 Proof of Lemma A.3

Proof. Define $\widehat{V}=\lim_{dt\to 0}V_{dt}$. Lemma A.3 is trivially true when μ is degenerate since $V(\mu)=\widehat{V}(\mu)=F(\mu)$. By induction, it is sufficient to prove that $\widehat{V}=V$ on the relative interior points of ΔX conditional on $\widehat{V}=V$ on $\partial\Delta X$ (the boundary of ΔX). The proof requires three steps. The first step shows that $\widehat{V}$ is unimprovable in an extended HJB equation. The second step shows that $\widehat{V}\geq V$. The last and most complicated step is to show that $V\geq\widehat{V}$.

Unimprovability: First, I define an extension of Equation (3) whose domain contains $\widehat{V}$. Let $\mathcal{L}$ denote the space of pointwise Lipschitz functions on $\Delta(X)$:

$$\mathcal{L}=\left\{V:\Delta(X)\mapsto\mathbb{R}^+\middle|\forall\mu\in\Delta X,\nu\in\Delta(\text{supp}(\mu)),\ \overline{\lim_{\nu\to\mu}}\frac{|V(\nu)-V(\mu)|}{\|\nu-\mu\|}\in\mathbb{R}\right\},$$

where $\|\cdot\|$ is the Euclidean distance on ΔX. A technical lemma S.5 guarantees that $\widehat{V}\in\mathcal{L}$.[30] Extend Equation (3) to $\mathcal{L}$:

$$\begin{aligned}\rho V(\mu)=\max\Bigg\{\rho F(\mu),\sup_{\substack{\nu_i\in\Delta(\text{supp}(\mu)),\\ p_i\in\mathbb{R}^+,\\ \widehat{\sigma}\in\mathbb{R}^{|\text{supp}(\mu)|-1}}}&\sum p_i(V(\nu_i)-V(\mu))-DV\left(\mu,\sum p_i\nu_i-\mu\right)\left\|\left(\sum p_i\nu_i-\mu\right)\right\|\\ &+\frac{1}{2}\|\widehat{\sigma}\|^2D^2V(\mu,\widehat{\sigma}) \qquad (16)\\ &-C\left(-\sum p_i(H(\nu_i)-H(\mu)-\nabla H(\mu)\cdot(\nu_i-\mu))-\frac{1}{2}\widehat{\sigma}^T\cdot\text{Hess}H(\mu)\cdot\widehat{\sigma}\right)\Bigg\},\end{aligned}$$

where D and D^2 replaces the Jacobian and Hessian operators, defined as follows.

Definition 3. *$\forall f\in\mathcal{L}$, $\forall x,y\in\Delta(X)$ and $\|y\|>0$,* $\begin{cases}Df(x,y)=\underline{\lim}_{\delta\to 0}\frac{f(x)-f(x-\delta y)}{\delta\|y\|}\\ D^2f(x,y)=\overline{\lim}_{\delta\to 0}2\frac{f(x+\delta y)-f(x)-\delta\cdot Df(x,y)}{\delta\|y\|^2}\end{cases}$.

Remark A.2. D denotes the lower Dini directional derivative. D^2 is the second-order analogy of D. Equation (16) is identical to Equation (3) on $C^1\Delta(X)$. Equation (16) is a slightly strengthened version of viscosity solution. Equation (16) uses test functions that are twice directional differentiable, which is a superset of globally C^2 test functions used in standard viscosity solutions.

I show that $\widehat{V}$ is unimprovable in Equation (16). Suppose for the sake of contradiction that $\widehat{V}$ is improvable at interior μ; then, there exists $p_i,\nu_i,\widehat{\sigma},I$ s.t.

$$\rho\widehat{V}(\mu)<\sum p_i\left(\widehat{V}(\nu_i)-\widehat{V}(\mu)\right)-D\widehat{V}\left(\mu,\sum p_i\nu_i-\mu\right)\left\|\sum p_i\nu_i-\mu\right\|+D^2\widehat{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2-C(I),$$

where $I=-\sum p_i(H(\nu_i)-H(\mu)-\nabla H(\mu)\cdot(\nu_i-\mu))-\widehat{\sigma}^T\text{Hess}H(\mu)\widehat{\sigma}$.

[30] The proof of Lemma S.5 uses the same technique as the following proof, hence is relegated to the supplemental materials.

Then, comparing the following two ratios,

$$\frac{\sum p_i\left(\widehat{V}(\nu_i)-\widehat{V}(\mu)\right)-D\widehat{V}(\mu,\sum p_i\nu_i-\mu)\|\sum p_i\nu_i-\mu\|}{-\sum p_i(H(\nu_i)-H(\mu)-\nabla H(\mu)\cdot(\nu_i-\mu))};\ \frac{D^2\widehat{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2}{-\widehat{\sigma}^T\mathrm{Hess}H(\mu)\widehat{\sigma}},$$

at least one of them must be larger than $\frac{\rho\widehat{V}(\mu)+C(I)}{I}$.

- *Case 1:*

$$\frac{\sum p_i\left(\widehat{V}(\nu_i)-\widehat{V}(\mu)\right)-D\widehat{V}(\mu,\sum p_i\nu_i-\mu)\|\sum p_i\nu_i-\mu\|}{-\sum p_i(H(\nu_i)-H(\mu)-\nabla H(\mu)\cdot(\nu_i-\mu))}>\frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}.$$

By Definition 3, there exists $\delta,\varepsilon>0$ s.t. :

$$\frac{\sum p_i\left(\widehat{V}(\nu_i)-\widehat{V}(\mu)\right)-\frac{\widehat{V}(\mu)-\widehat{V}(\mu-\delta(\sum p_i\nu_i-\mu))}{\delta}}{\sum p_i(H(\mu)-H(\nu_i))+\frac{H(\mu)-H(\mu-\delta(\sum p_i\nu_i-\mu))}{\delta}}\geq\frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}+\varepsilon, \tag{17}$$

where δ is sufficiently small that $\nu_0=\mu-\delta(\sum p_i\nu_i-\mu)\in\Delta X^{o}$. Define $p_0'=\frac{1}{1+\delta}$ and $p_i'=\frac{\delta}{1+\delta}p_i$. Then, (p_i',ν_i) is a probability measure ($\sum p_i'=1$) and satisfies Bayesian plausible ($\sum p_i'\nu_i=\mu$), where 0 is also included in indices i's. Replacing terms in Equation (17) and letting $I'=H(\mu)-\sum p_i'H(\nu_i)$:

$$\frac{\sum p_i'\widehat{V}(\nu_i)-\widehat{V}(\mu)}{-\sum p_i'H(\nu_i)+H(\mu)}\geq\frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}+\varepsilon$$
$$\Longrightarrow\sum p_i'\widehat{V}(\nu_i)-\frac{I'}{I}C(I)\geq\left(1+\rho\frac{I'}{I}\right)\widehat{V}(\mu)+\varepsilon I'. \tag{18}$$

It is easy to verify that I' is continuous in δ and zero when $\delta=0$. Thus, δ can be chosen sufficiently small s.t.

$$e^{\rho\frac{I'}{I}}-\left(1+\rho\frac{I'}{I}\right)=\sum_{k=1}^{\infty}\frac{1}{(k+1)!}\left(\frac{\rho}{I}\right)^{k+1}I(\nu_i|\mu)^k\cdot I'\leq\frac{\varepsilon I'}{2\sup F}. \tag{19}$$

The equality stems from the Taylor expansion of e^x. Plug Equation (19) into Equation (18):

$$\sum p_i'\widehat{V}(\nu_i)-\frac{I'}{I}C(I)\geq e^{\rho\frac{I'}{I}}\widehat{V}(\mu)+\frac{\varepsilon}{2}I'$$
$$\Longrightarrow e^{-\rho\frac{I'}{I}}\left(\sum p_i'\widehat{V}(\nu_i)\right)-\frac{I'}{I}C(I)\geq\widehat{V}(\mu)+e^{-\rho\frac{I'}{I}}\frac{\varepsilon I'}{2}-\left(1-e^{-\rho\frac{I'}{I}}\right)\frac{I'}{I}C(I). \tag{20}$$

Note that $\left(1-e^{-\rho\frac{I'}{I}}\right)I'$ is a second-order small term. Then, select δ such that Equation (20) implies

$$e^{-\rho\frac{I'}{I}}\left(\sum p_i'\widehat{V}(\nu_i)\right)-\frac{I'}{I}C(I)\geq\widehat{V}(\mu)+\frac{\varepsilon}{4}I'.$$

Henceforth, fix ε and δ. Select $dt=\frac{I'}{I}$, $dt_m=\frac{dt}{m}$. By the uniform convergence of V_{dt}, there exists N s.t. $\forall m\geq N$:

$$e^{-\rho dt}\left(\sum p_i'V_{dt_m}(\nu_i)\right)-dt_m\cdot C\left(\frac{I'/m}{dt_m}\right)>V_{dt_m}(\mu)$$

$$\Longrightarrow e^{-\rho m dt_m}\left(\sum p_i' V_{dt_m}(\nu_i)\right)-\sum_{\tau=0}^{m-1} e^{-\rho\tau dt_m} C_{dt_m}\left(\frac{I'}{m}\right) > V_{dt_m}(\mu).$$

That is, in the dt_m-discrete-time problem, an admissible strategy that divides the experiment (p_i',ν_i) into m periods (based on Lemma S.2 in Section S1.1 of the supplemental material) and follows the optimal strategy of V_{dt_m} at the end of the m periods strictly improves V_{dt_m}. Contradiction.

- *Case 2*:

$$\frac{D^2\widehat{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2}{-\widehat{\sigma}^T \mathrm{Hess}H(\mu)\widehat{\sigma}} > \frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}.$$

Then, by Definition 3, there exists $\widehat{\sigma}$, $\delta,\varepsilon>0$ s.t.:

$$\frac{\widehat{V}(\mu+\delta\widehat{\sigma})-\widehat{V}(\mu)-\delta D\widehat{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|}{-H(\mu+\delta\widehat{\sigma})+H(\mu)+\delta\nabla H(\mu)\cdot\widehat{\sigma}} \geq \frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}+2\varepsilon.$$

Then, by the definition of operator D in Definition 3, there exists δ' s.t.

$$\frac{\widehat{V}(\mu+\delta\widehat{\sigma})-\widehat{V}(\mu)-\delta\frac{\widehat{V}(\mu)-\widehat{V}(\mu-\delta'\widehat{\sigma})}{\delta'}}{-H(\mu+\delta\widehat{\sigma})+H(\mu)+\delta\frac{H(\mu)-H(\mu-\delta'\widehat{\sigma})}{\delta'}} \geq \frac{\rho}{I}\widehat{V}(\mu)+\frac{C(I)}{I}+\varepsilon.$$

Let $\nu_1=\mu-\delta'\widehat{\delta}$, $\nu_2=\mu+\delta\widehat{\sigma}$, $p_1=\frac{\delta'}{\delta+\delta'}, p_2=\frac{\delta}{\delta+\delta'}$ and $I'=H(\mu)-\sum p_i H(\nu_i)$. Then,

$$\sum p_i\widehat{V}(\nu_i)\geq\left(1+\rho\frac{I'}{I}\right)\widehat{V}(\mu)+\frac{I'}{I}C(I)+\varepsilon I'. \tag{21}$$

Note that Equation (21) is exactly the same as Equation (18) in Case 1. An identical argument rules out this case as well.

Equality: I show that $\forall$ C^1 functions V solving Equation (16), $\widehat{V}=V$. I prove the inequality from both directions for $\mu\in\Delta(X)^{\mathrm{o}}$:

- $\widehat{V}(\mu)\geq V(\mu)$: Suppose not. Then, consider $U(\mu)=\widehat{V}(\mu)-V(\mu)$. Since both V and $\widehat{V}$ are continuous, U is continuous. Therefore, $\arg\min U$ is non-empty, and $\min U<0$ according to our assumption. Choose $\mu\in\arg\min U$ ($\mu\in\Delta X^o$ since $V=\widehat{V}$ on the boundary). $\widehat{V}(\mu)\geq F(\mu)$ implies that $V(\mu)>F(\mu)$. Let $(p,\nu,\hat{\sigma})$ be a strategy that achieves $V(\mu)$ at μ in Equation (3):

$$\begin{aligned}\rho V(\mu)=&p(V(\nu)-V(\mu)-\nabla V(\mu)(\nu-\mu))+\frac{1}{2}D^2V(\mu,\widehat{\sigma})\|\hat{\sigma}\|^2\\&-C\left(-p(H(\nu)-H(\mu)-\nabla H(\mu)(\nu-\mu))-\frac{1}{2}\widehat{\sigma}^T\mathrm{Hess}H(\mu)\widehat{\sigma}\right).\end{aligned} \tag{22}$$

Now compare $D\widehat{V}$ and DV:

$$\frac{\widehat{V}(\mu)-\widehat{V}(\mu')}{\|\mu-\mu'\|}=\frac{V(\mu)-V(\mu')+U(\mu)-U(\mu')}{\|\mu-\mu'\|}\leq\frac{V(\mu)-V(\mu')}{\|\mu-\mu'\|}$$

$$\implies \lim_{\mu'\to\mu} \frac{\widehat{V}(\mu)-\widehat{V}(\mu')}{\|\mu-\mu'\|} \leq \lim_{\mu'\to\mu} \frac{V(\mu)-V(\mu')}{\|\mu-\mu'\|}$$
$$\implies D\widehat{V}(\mu,\mu'-\mu)\|\mu'-\mu\| \leq \nabla V(\mu)\cdot(\mu'-\mu).$$

Compare $D^2\widehat{V}$ and D^2V:

$$\frac{\widehat{V}(\mu')-\widehat{V}(\mu)-D\widehat{V}(\mu,\mu'-\mu)\|\mu'-\mu\|}{\|\mu'-\mu\|^2} \geq \frac{V(\mu')-V(\mu)-\nabla V(\mu)\cdot(\mu'-\mu)+U(\mu')-U(\mu)}{\|\mu-\mu'\|^2}$$
$$\implies D^2\widehat{V}(\mu,\widehat{\sigma}) \geq D^2V(\mu,\widehat{\sigma}).$$

Therefore, Equation (22) implies

$$\begin{aligned}\rho V(\mu) \leq\, & p\Big(\widehat{V}(\nu)-\widehat{V}(\mu)-(U(\nu)-U(\mu))\Big)\\ &-D\widehat{V}(\mu,\nu-\mu)\|\nu-\mu\|+\frac{1}{2}D^2\widehat{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2\\ &-C\left(-p(H(\nu)-H(\mu)+\nabla H(\mu)(\nu-\mu))-\frac{1}{2}\widehat{\sigma}^T\text{Hess}H(\mu)\widehat{\sigma}\right)\\ \leq\, & \rho\widehat{V}(\mu).\end{aligned}$$

The first inequality is from replacing DV and D^2V with $D\widehat{V}$ and $D^2\widehat{V}$. The second inequality is from $U(\nu)-U(\mu)\geq 0$ and the unimprovability of $\widehat{V}$. Contradiction.

- $V(\mu)\geq\widehat{V}(\mu)$: I prove via a contradiction that, suppose $V(\mu)<\lim V_{dt}(\mu)$, then the strategies that achieve $V_{dt}(\mu)$s converge to a continuous-time strategy that strictly improves $V(\mu)$ and violates unimprovability. The technical difficulty in this step is an accounting exercise that tracks the approximation error in the limit.

  Suppose $V(\mu')>\lim V_{dt}(\mu')$; there exists dt s.t. $V_{dt}(\mu')>V(\mu')$. Let $dt_n=\frac{dt}{2^n}$. Since V_{dt_n} increases in n, there exists $\varepsilon>0$ s.t. $V_{dt_n}(\mu')-V(\mu')\geq\varepsilon$ $\forall n\in\mathbb{N}$. Now consider $U_n=V-V_{dt_n}$. U_n is continuous by Lemma A.2 and $U_n(\mu')\leq-\varepsilon$. Select $\mu^n\in\arg\min U_n$. Since $\Delta(X)$ is compact, μ^n can be chosen without loss such that $\lim\mu^n=\mu$. $U_n(\mu^n)\leq U_n(\mu')\leq-\varepsilon$; therefore, since $U(\mu)=\lim U_n(\mu^n)\leq-\varepsilon$, μ must be in the interior of $\Delta(X)$. Thus, μ^n can be selected without loss such that $\mu^n\in\Delta(X)^o$. For each n, let $\big((p_i^n,\nu_i^n),I_n\big)$ be the optimal strategy of the discrete-time problem.

$$\begin{aligned}\sum p_i^n(V(\nu_i^n)-V(\mu^n)) &= \sum p_i^n(V_{dt_n}(\nu_i^n)-V_{dt_n}(\mu^n)-U_n(\mu^n)+U(\nu_i^n))\\ &\geq \sum p_i^n(V_{dt_n}(\nu_i^n)-V_{dt_n}(\mu^n))\\ &= \left(e^{\rho dt_n}-1\right)V_{dt_n}(\mu^n)+e^{\rho dt_n}dt_nC(I_n)\\ &\geq \rho dt_nV_{dt_n}(\mu^n)+e^{\rho dt_n}dt_nC(I_n)\\ &\geq \rho dt_n\varepsilon+\rho dt_nV(\mu^n)+e^{\rho dt_n}dt_nC(I_n)\end{aligned}$$
$$\implies \rho V(\mu^n)\leq-\rho\varepsilon+\sum\frac{p_i^n}{dt_n}(V(\nu_i^n)-V(\mu^n))-e^{\rho dt_n}C(I_n)$$

$$\Longrightarrow \rho V(\mu^n) \leq -\rho\varepsilon + \sum \frac{p_i^n}{dt_n}(V(\nu_i^n) - V(\mu^n)) - C(I_n). \tag{23}$$

The first equality is by the definition of U_n. The first inequality is from $\mu^n \in \arg\min U_n$. The second inequality is from $e^x - 1 \geq x$. The third inequality is from $U_n(\mu^n) \leq -\varepsilon$. Therefore, $\rho V(\mu^n) + \rho\varepsilon$ is a *lower bound* for the payoff from strategy $((p_i^n, \nu_i^n), I_n)$.

Since the number of posteriors ν_i^n is no more than $2|X|$, a subsequence of ν_i^n converges. WLOG, let $\lim_n \nu_i^n = \nu_i$. Partition the ν_i^ns into two subsets: $\lim \nu_i^n = \nu_i \neq \mu$ and $\lim \nu_j^n = \mu$. Given $\sum p_{i,j}^n(\nu_{i,j}^n - \mu_n) = 0$, insert $\sum p_{i,j}^n \nabla V(\mu^n)(\nu_{i,j}^n - \mu^n)$ into $\sum p_{i,j}^n(V(\nu_{i,j}^n) - V(\mu^n))$. Then, the ν_i^ns approximate Poisson signals, and the ν_j^ns approximate Gaussian signals. In what follows, I utilize Equation (16) to derive an *upper bound* for the payoff from strategy $((\nu_i^n, p_i^n), I^n)$. Since V is unimprovable, $\forall c, \widehat{\sigma}$ and I, $D^2V(\mu, \widehat{\sigma})\|\widehat{\sigma}\|^2 \leq -\widehat{\sigma}^T \mathrm{Hess} H(\mu)\widehat{\sigma}\left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)$. Since $V \in C^1$, $H \in C^2$, $\forall \eta$, $\exists \delta$ s.t. $\forall |\mu' - \mu| \leq \delta$:

$$\begin{cases} \|\mathrm{Hess} H(\mu) - \mathrm{Hess} H(\mu')\| \leq \eta \\ |V(\mu) - V(\mu')| \leq \eta \end{cases}$$

$$\Longrightarrow D^2V(\mu', \widehat{\sigma}) \leq \left(\frac{\rho}{I}V(\mu') + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{Hess} H(\mu')\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right)$$
$$\leq \left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{Hess} H(\mu)\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right) + \left(\frac{\rho}{I}\sup F + \frac{C(I)}{I}\right)\eta + \frac{\rho}{I}\eta\|\mathrm{Hess} H(\mu)\|.$$

If η and δ are selected properly,

$$D^2V(\mu', \widehat{\sigma}) \leq \left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{Hess} H(\mu)\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right) + \frac{1 + C(I)}{I}\eta.$$

Since $\lim \nu_j^n = \lim \mu^n = \mu$, there exists N s.t. $\forall n \geq N$, $\left|\nu_j^n - \mu\right| < \delta$, and $|\mu^n - \mu| < \delta$. Now I want to obtain a second-order approximation of $V(\nu_j^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_j^n - \mu^n)$. To apply Taylor expansion to a not necessarily twice-differentiable function V, I apply an extended mean value theorem (Lemma S.6, proof relegated to the supplemental materials) to $g(\alpha) = V(\alpha\nu_j^n + (1-\alpha)\mu^n)$:

$$\begin{aligned}
&V(\nu_j^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_j^n - \mu^n) = g(1) - g(0) - g'(0) \\
\leq &\frac{1}{2}\sup_{\alpha \in (0,1)} D^2 g(\alpha, 1) = \sup_{\alpha \in (0,1)} \overline{\lim_{d \to 0}} \frac{g(\alpha + d) - g(\alpha) - g'(\alpha)d}{d^2} \\
= &\sup_{\xi \in (\mu^n, \nu_j^n)} \overline{\lim_{d \to 0}} \frac{V(\xi + d(\nu_j^n) - \mu^n) - V(\xi) - dJV(\xi)(\nu_j^n - \mu^n)}{d^2} \\
\leq &\frac{1}{2}\sup_{|\xi - \mu| \leq \delta} D^2 V(\xi, \nu_j^n - \mu^n)\left\|\nu_j^n - \mu^n\right\|^2
\end{aligned}$$

$$\leq -\frac{1}{2}\left(\frac{\rho}{I}V(\mu)+\frac{C(I)}{I}\right)(\nu_j^n-\mu^n)^T \text{Hess} H(\mu)(\nu_j^n-\mu^n)+\frac{1+C(I)}{2I}\eta\left\|\nu_j^n-\mu^n\right\|^2. \quad (24)$$

Therefore, by applying Equation (24),

$$\begin{aligned}
&\sum p_{i,j}^n\left(V(\nu_{i,j}^n)-V(\mu^n)\right)\\
=&\sum p_i^n(V(\nu_i^n)-V(\mu^n)-\nabla V(\mu^n)(\nu_i^n-\mu^n))+\sum p_j^n\left(V(\nu_j^n)-V(\mu^n)-\nabla V(\mu^n)\left(\nu_j^n-\mu^n\right)\right)\\
\leq&\sum p_i^n(V(\nu_i^n)-V(\mu^n)-\nabla V(\mu^n)(\nu_i^n-\mu^n))\\
&-\frac{1}{2}\left(\frac{\rho}{I}V(\mu)+\frac{C(I)}{I}\right)\sum p_j^n\left(\nu_j^n-\mu^n)^T\text{Hess}H(\mu)(\nu_j^n-\mu^n)\right)+\frac{1+C(I)}{2I}\eta\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2.
\end{aligned} \quad (25)$$

Note that Equations (24) and (25) hold for all Is. Now let $\bar{p}_i^n=\frac{p_i^n}{dt_n}$ and $-\widehat{\sigma}_n^T\text{Hess}H(\mu^n)\widehat{\sigma}^n dt_n=\sum p_j^n\left(H(\mu^n)-H(\nu_j^n)+\nabla H(\mu)(\nu_j^n-\mu^n)\right)$. Thus,

$$\sum\bar{p}_i^n(H(\mu^n)-H(\nu_i^n)+H'(\mu^n)(\nu_i^n-\mu^n))-\widehat{\sigma}_n^T\text{Hess}H(\mu^n)\widehat{\sigma}_n=I_n. \quad (26)$$

$(\bar{p}_i^n,\nu_i^n,\widehat{\sigma}_n)$ is a feasible experiment for Equation (16). Therefore, by the optimality of V at μ^n,

$$\begin{cases}\sum\bar{p}_i^n(V(\nu_i^n)-V(\mu^n)-\nabla V(\mu^n)(\nu_i^n-\mu^n))\leq\left(I_n+\widehat{\sigma}^{nT}\text{Hess}H(\mu^n)\widehat{\sigma}^n\right)\left(\frac{\rho}{I_n}V(\mu^n)+\frac{C(I_n)}{I_n}\right)\\
D^2V(\mu^n,\widehat{\sigma}^n)\leq-\frac{\widehat{\sigma}^{nT}\text{Hess}H(\mu)\widehat{\sigma}^n}{\|\widehat{\sigma}^n\|^2}\left(\frac{\rho}{I_n}\widehat{V}(\mu^n)+\frac{C(I_n)}{I_n}\right).\end{cases} \quad (27)$$

Then, study the term $\sum p_j^n(\nu_j^n-\mu^n)^2$. Approximate $g(\alpha)=H(\alpha\nu_j^n+(1-\alpha)\mu^n)$ in the second order by applying Lemma S.6 again:

$$\begin{aligned}
&\sum p_j^n\left(H(\mu^n)-H(\nu_j^n)+\nabla H(\mu^n)(\nu_j^n-\mu^n)\right)\\
\geq&\frac{1}{2}\inf_{\xi_j^n\in[\mu^n,\nu_n^j]}\sum p_j^n\left(-(\nu_j^n-\mu^n)^T\text{Hess}H(\xi_j^n)(\nu_j^n-\mu^n)\right)\\
\geq&-\frac{1}{2}\sum p_j^n((\nu_j^n-\mu^n)^T\text{Hess}H(\mu)(\nu_j^n-\mu^n))-\frac{1}{2}\eta\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2.
\end{aligned} \quad (28)$$

Therefore,

$$\begin{aligned}
\sum\frac{p_{i,j}^n}{dt_n}\left(V(\nu_{i,j}^n)-V(\mu^n)\right)\leq&\sum\bar{p}_i^n(V(\nu_i^n)-V(\mu^n)-\nabla V(\mu^n)(\nu_i^n-\mu^n))\\
&+\frac{1}{2}\sum\frac{p_j^n}{dt_n}\left(-(\nu_j^n-\mu^n)^T\text{Hess}H(\mu)(\nu_j^n-\mu^n)\left(\frac{\rho}{I_n}V(\mu)+\frac{C(I_n)}{I_n}\right)\right)\\
&+\sum\frac{p_j^n}{dt_n}\left(\frac{1+C(I_n)}{2I_n}\eta\left\|\nu_j^n-\mu^n\right\|^2\right)\\
\leq&\left(I_n+\widehat{\sigma}^{nT}\text{Hess}H(\mu^n)\widehat{\sigma}^n\right)\left(\frac{\rho}{I_n}V(\mu^n)+\frac{C(I_n)}{I_n}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\left(\sum \frac{p_j^n}{dt_n}(H(\mu^n)-H(\nu_j^n)+\nabla H(\mu^n)(\nu_j^n-\mu^n))\right.\\
&\left.+\frac{1}{dt_n}\frac{\eta}{2}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\right)\left(\frac{\rho}{I_n}V(\mu)+\frac{C(I_n)}{I_n}\right)\\
&+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\frac{1+C(I_n)}{2I_n}\eta\\
=&\left(I_n+\widehat{\sigma}^{nT}\mathrm{Hess}H(\mu^n)\widehat{\sigma}^n\right)\left(\frac{\rho}{I_n}V(\mu^n)+\frac{C(I_n)}{I_n}\right)\\
&+\left(-\widehat{\sigma}^{nT}\mathrm{Hess}H(\mu^n)\widehat{\sigma}^n+\frac{1}{dt_n}\frac{\eta}{2}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\right)\left(\frac{\rho}{I_n}V(\mu)+\frac{C(I_n)}{I_n}\right)\\
&+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\frac{1+C(I_n)}{2I_n}\eta\\
\le&\rho V(\mu^n)+C(I_n)+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\left(\frac{1+\rho V(\mu)+2C(I_n)}{2I_n}\right)\eta+\rho\eta.
\end{aligned}
$$

The first inequality is Equation (25). The second inequality stems from Equation (27) and Equation (28). The next equality stems from the definition of $\widehat{\sigma}_n^2$. The last inequality stems from canceling out terms and $-\widehat{\sigma}^{nT}\mathrm{Hess}H(\mu^n)\widehat{\sigma}^n\le I_n$ (note the difference between $V(\mu)$ and $V(\mu^n)$). Then, by plugging into Equation (23):

$$
\rho V(\mu^n)\le-\rho\varepsilon+\rho V(\mu^n)+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\left(\frac{1+\rho V(\mu)+2C(I_n)}{2I_n}\right)\eta+\rho\eta.
$$

Moreover,

$$
\begin{aligned}
&\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\inf_\sigma\frac{\left|\sigma^T\mathrm{Hess}H(\mu)\sigma\right|}{\|\sigma\|^2}\\
&\le\sum p_j^n(\nu_j^n-\mu^n)^T\mathrm{Hess}H(\mu)(\mu^n-\mu^n)\le I_n dt_n+\eta\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\\
\Longrightarrow&\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\le\frac{I_n dt_n}{\inf_\sigma\frac{\left|\sigma^T\mathrm{Hess}H(\mu)\sigma\right|}{\|\sigma\|^2}-\eta}\\
\Longrightarrow&\rho\varepsilon\le\frac{1}{2}(1+\rho V(\mu)+2C(I_n))\frac{\eta}{\inf_\sigma\frac{\left|\sigma^T\mathrm{Hess}H(\mu)\sigma\right|}{\|\sigma\|^2}-\eta}+\rho\eta.
\end{aligned}
$$

Lemma A.4 below proves that $\{C(I_n)\}$ are uniformly bounded. Since H is strictly concave, $\inf_\sigma\frac{\left|\sigma^T\mathrm{Hess}H(\mu)\sigma\right|}{\|\sigma\|^2}$ is positive. The inequality holds when η is chosen to be smaller than $\inf_\sigma\frac{\left|\sigma^T\mathrm{Hess}H(\mu)\sigma\right|}{\|\sigma\|^2}$. By taking $\eta\to0$, the LHS is eventually larger than the RHS, which is a contradiction. Therefore,

$$
V(\mu)=\overline{\lim_{dt\to0}}V_{dt}(\mu)=\widehat{V}(\mu). \qquad \square
$$

Lemma A.4. *Given Assumption 2, there $\exists\Delta\in\mathbb{R}^+$ s.t. $\forall dt$, suppose (p_i,ν_i) is the optimal policy at μ in Equation (13), then $\sum p_i(H(\mu)-H(\nu_i))\le\Delta dt$.*

Proof. $\forall dt>0$, $\forall \mu\in(0,1)$, let (p_i,ν_i) be the optimal policy at μ in Equation (13) and $I=\frac{\sum p_i(H(\mu)-H(\nu_i))}{dt}$. The optimality of the strategy implies:

$$e^{-\rho dt}\sum_{i=1}^{2|X|}p_iV_{dt}(\nu_i)-C(I)dt\geq e^{-\rho dt}\left(\sum_{i=1}^{2|X|}\frac{p_i}{2}V_{dt}(\nu_i)+\frac{1}{2}V_{dt}(\mu)\right)-C\left(\frac{I}{2}\right)dt. \tag{29}$$

The RHS of Equation (29) is the payoff from a strategy that mixes (p_i,ν_i) with trivial information with $\frac{1}{2}$ probabiltiy. Equations (13) and (29) implies

$$\begin{aligned}&\frac{1}{2}e^{-\rho dt}\sum_{i=1}^{2|X|}p_iV_{dt}(\nu_i)-C(I)dt\geq\frac{1}{2}e^{-2\rho dt}\sum_{i=1}^{2|X|}p_iV_{dt}(\nu_i)-\frac{1}{2}e^{-\rho dt}C(I)dt-C\left(\frac{I}{2}\right)dt\\ \Longrightarrow&\left(1-\frac{1}{2}e^{-\rho dt}\right)C(I)-C\left(\frac{I}{2}\right)\leq\frac{e^{-\rho dt}-e^{-2\rho dt}}{2dt}\sum_{i=1}^{2|X|}p_iV_{dt}(\nu_i)\\ \Longrightarrow&\frac{1}{2}C(I)-C\left(\frac{I}{2}\right)\leq\frac{\rho}{2}\sup F.\end{aligned}$$

The last inequality is from $e^{-\rho dt}<1$ and $1-e^{-x}<x$. Suppose for the purpose of contradiction that I is unbounded. This implies $\lim_{x\to\infty}\frac{1}{2}C(x)-C(\frac{1}{2}x)\leq\frac{\rho\sup F}{2}$; hence, $\exists M$ s.t. $\frac{1}{2}C(x)-C(\frac{1}{2}x)\leq M$. Then, $\forall n\in\mathbb{N}$, $C(2^n)+M\leq 2\big(C(2^{n-1})+M\big)\leq 2^n(C(1)+M)$. This means $C'(I)$ is bounded and contradicts $\lim_{I\to\infty}C'(I)=\infty$ in Assumption 2. □

B Proof of Theorem 2

Proof. I prove Theorem 2 by guess and verification. To simplify notations, let $J(\mu,\nu)=H(\mu)-H(\nu)+H'(\mu)(\nu-\mu)$ (the Bregman divergence associated with $-H$).

Construction of $V(\mu)$:

- *Step 1*: Find the critical belief μ^*. Define:

$$\begin{cases}\overline{V}^+(\mu)=\max\limits_{\nu\geq\mu,I}\dfrac{F(\nu)-\frac{C(I)}{I}J(\mu,\nu)}{1+\frac{\rho}{I}J(\mu,\nu)}\\ \overline{V}^-(\mu)=\max\limits_{\nu\leq\mu,I}\dfrac{F(\nu)-\frac{C(I)}{I}J(\mu,\nu)}{1+\frac{\rho}{I}J(\mu,\nu)}\end{cases}. \tag{30}$$

  Evidently, $\overline{V}^+\geq F$ and $\overline{V}^-\geq F$. Lemma B.1 proves that $I^+=\{\mu|\overline{V}^+(\mu)>F(\mu)\}$ is an open interval containing $\{\mu|F'(\mu)>0\}$ and $I^-=\{\mu|\overline{V}^-(\mu)>F(\mu)\}$ is an open interval containing $\{\mu|F'(\mu)<0\}$. When $I^+\cap I^-\neq\emptyset$, there exists unique interior μ^* s.t. $\overline{V}^+(\mu^*)=\overline{V}^-(\mu^*)>F(\mu^*)$. When $I^+(I^-)=\emptyset$, define $\mu^*=1(0)$. When I^+, I^- are both non-empty but do not intersect, there exists $\mu^*\in[\sup I^-,\inf I^+]$ and $F'(\mu^*)=0$. [31]

- *Step 2:* I construct the first piece of $V(\mu)$ to the right of μ^*. There are two possible

[31] If H and F are symmetric, then $\overline{V}^+$ and $\overline{V}^-$ are mirror images of each other and $\mu^*=0.5$.

cases to be discussed. [32]

Case 1: Suppose $I^+ \cap I^- \neq \emptyset$. Then, the initial conditions $\mu_0 = \mu^*, V_0 = \overline{V}(\mu^*), V_0' = 0$ satisfy the conditions of Lemma B.3, which states that there exists $\mu^{**} > \mu^*$ and C^1 strictly increasing function $V_{\mu^*}(\mu) > F(\mu)$ defined on $[\mu^*, \mu^{**})$ satisfying

$$V_{\mu^*}(\mu) = \max_{\nu \geq \mu, I} \frac{I}{\rho} \frac{F(\nu) - V_{\mu^*}(\mu) - V_{\mu^*}'(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho}. \tag{31}$$

Equation (31) has a maximizer $\nu(\mu)$ that is piecewise unique, C^1, and strictly decreasing. Moreover, when $\mu \to \mu^{**}-$, $V_{\mu^*}(\mu) \to F(\mu^{**})$ and $V_{\mu^*}'(\mu) \to F'(\mu^{**})$. Evidently, the optimal $I(\mu)$ satisfies $I(\mu)C'(I(\mu)) - C(I(\mu)) = \rho V_{\mu^*}(\mu)$ and strictly increases with $V_{\mu^*}(\mu)$.

Case 2: Suppose $I^+ \neq \emptyset$ and $I^+ \cap I^- = \emptyset$; let $\mu^{**} = \mu^*$.

- *Step 3:* Construct V to the right of μ^{**}. Define

$$U(\mu) = \sup_{\nu \geq \mu, I} \frac{I}{\rho} \frac{F(\nu) - F(\mu) - F'^+(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho} \tag{32}$$

$U(\mu)$ denotes the flow payoff from choosing the optimal Poisson signal with $\nu \geq \mu$ given $F(\mu)$ as the hypothetical value function. $U(\mu)$ is continuous in the linear regions of F and jumps down at the kinks of F; hence, U is lower semi-continuous and right continuous. Let $\Gamma = \{\mu > \mu^{**} | U(\mu) > F(\mu)\}$. Lemma B.2 implies that $\inf \Gamma > 0$. The definition of μ^{**} guarantees that $U(\mu^{**}) \leq F(\mu^{**})$.[33] Since U is lower semi-continuous, Γ is an open set; hence, Γ is a countable union of open intervals. Let Ω be the set of the left boundaries of these intervals. Since U is right continuous, $\forall \mu^\diamond \in \Omega$, $U(\mu^\diamond) = F(\mu^\diamond)$ and the initial conditions $\mu^\diamond, F(\mu^\diamond), F'^+(\mu^\diamond)$ satisfy the conditions of Lemma B.3. There exists $\mu^{\diamond\diamond} > \mu^\diamond$ and $V_{\mu^\diamond}$ defined on $[\mu^\diamond, \mu^{\diamond\diamond})$ s.t.

$$V_{\mu^\diamond}(\mu) = \max_{\nu \geq \mu^\diamond, I} \frac{I}{\rho} \frac{F(\nu) - V_{\mu^\diamond}(\mu) - V_{\mu^\diamond}'(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho}. \tag{31}$$

Extend the definition of $V_{\mu^\diamond}$ to $V_{\mu^\diamond}(\mu) = F(\mu)$ when $\mu \notin [\mu^\diamond, \mu^{\diamond\diamond})$. Define:

$$V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ \sup\limits_{\mu^\diamond \in \Omega} \{V_{\mu^\diamond}(\mu)\} & \text{if } \mu \geq \mu^{**} \end{cases}. \tag{33}$$

On the left of μ^*, V is constructed using a symmetric argument.

Smoothness:

I claim that $V(\mu)$ defined as Equation (33) is a C^1 function on $(0,1)$. This claim is

[32] Case one is the only conceptually non-trivial case. I omit the third case $\mu^* = 1$ because it is symmetric to the case $\mu^* = 0$.

[33] This inequality is straightforward when $\mu^{**} > \mu^*$ or $\mu^{**} = 0$. Otherwise, $F'(\mu^{**}) = 0$; hence $U(\mu^{**}) = \overline{V}^+(\mu^{**}) = F(\mu^{**})$.

purely for technical use. I relegate these technical proofs to Section S2.2. It shows (based on a standard no-crossing argument from the ODE theory) that on each interval that $V(\mu)>F(\mu)$, $\exists\mu^\diamond\in\Omega$ s.t. $V(\mu)\equiv V_{\mu^\diamond}(\mu)$.

Unimprovability:

- *Step 1*: Let $\Theta(\mu)=[\mu,1]$ when $\mu\ge\mu^*$ and $[0,\mu]$ when $\mu<\mu^*$. First, verify that
$$V(\mu)=\max\left\{F(\mu),\max_{\nu\in\Theta(\mu),I}\frac{I}{\rho}\frac{F(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}\right\}. \tag{34}$$
Equation (34) reduces to Equation (31) when $V(\mu)>F(\mu)$. Thus, it remains to prove Equation (34) when $V(\mu)=F(\mu)$. Let (a,b) be an interval that constitutes Γ. I claim that $V_a(\mu)>F(\mu)$ on (a,b). If not, then $\exists\mu\in(a,b)$ s.t. $V_a(\mu)\le F(\mu)$ and $V_a'(\mu)\le F'(\mu)$, which implies $V_a(\mu)\ge U(\mu)>F(\mu)$. Contradiction. Therefore, $V(\mu)>F(\mu)$ on Γ and Equation (34) is proved.

- *Step 2*: Verify that taking an immediate action following the posterior is optimal:
$$V(\mu)=\max\left\{F(\mu),\max_{\nu\in\Theta(\mu),I}\frac{I}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}\right\}. \tag{35}$$
The proof is explained visually in Section 4.2.1, part 3. Here, I formalize the argument. For $\mu\in E$ and $\mu\ge\mu^*$, the maximizer of Equation (34) is characterized by:
$$\begin{cases}\nu^*\in\arg\max_{\nu\ge\mu}\frac{F(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}\\ C'(I^*)=\max_{\nu\ge\mu}\frac{F(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}\end{cases}, \tag{36}$$
where $V(\mu)=\frac{I^*C'(I^*)-C(I^*)}{\rho}$. Suppose Equation (35) is violated:
$$C'(I^*)=\max_{\nu\ge\mu}\frac{F(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}<\sup_{\nu\ge\mu}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}.$$
Let $L(\nu)$ be defined as:
$$L(\nu)=V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)-C'(I^*)J(\mu,\nu)$$
By construction, $L(\mu)=0$, $L(1)\le0$, and $\sup_{\nu\le\mu}L(\nu)>0$. By continuity, L has a maximizer $\mu'\in(\mu,1)$. The FOC at μ' implies:
$$\frac{V'(\mu')-V'(\mu)}{H'(\mu)-H'(\mu')}=C'(I^*). \tag{37}$$
$L(\mu')>0$ implies:
$$\frac{V(\mu')-V(\mu)-V'(\mu)(\mu'-\mu)}{J(\mu,\mu')}>C'(I^*). \tag{38}$$
Note that $V(\mu')>F(\mu')$. Therefore, there exists ν',I' s.t. $\rho V(\mu')=I'C'(I')-C(I')$ and
$$C'(I')=\frac{F(\nu')-V(\mu')-V'(\mu')(\nu'-\mu')}{J(\mu',\nu')}. \tag{39}$$

Since $\mu'>\mu$, $V(\mu')>V(\mu)$ and hence $C'(I')>C'(I^*)$. Therefore,

$$\frac{F(\nu')-V(\mu)-V'(\mu)(\nu'-\mu)}{J(\mu,\nu')}$$
$$=\frac{\underbrace{\left[F(\nu')-V(\mu')-V'(\mu')(\nu'-\mu')\right]}_{A1}+\underbrace{\left[V(\mu')-V(\mu)-V'(\mu)(\mu'-\mu)\right]}_{A2}+\underbrace{\left[(V'(\mu')-V'(\mu))(\nu'-\mu')\right]}_{A3}}{\underbrace{\left[H(\mu')-H(\nu')+H'(\mu')(\nu'-\mu')\right]}_{B1}+\underbrace{\left[H(\mu)-H(\mu')+H'(\mu)(\mu'-\mu)\right]}_{B2}+\underbrace{\left[(H'(\mu)-H'(\mu'))(\nu'-\mu')\right]}_{B3}}$$
$$>C'(I^*).$$

The strict inequality is from that fact that $A_i,B_i>0$, $\frac{A_1}{B_1}=C'(I')>C'(I^*)$ (Equation (39)), $\frac{A_2}{B_2}>C'(I^*)$ (Equation (38)), and $\frac{A_3}{B_3}=C'(I^*)$ (Equation (37)). This contradicts Equation (36). A symmetric argument shows that Equation (35) is satisfied for $\mu<\mu^*$.

- *Step 3*: I show that V satisfies Equation (3), which is less restrictive than Equation (35) by allowing 1) Gaussian signals. 2) all possible posteriors instead of just confirmatory posteriors. First, since V is C^1, increasing and has a piecewise differentiable optimizer ν, the envelope theorem implies that $\forall\mu$ s.t. ν is differentiable:

$$\begin{aligned}
V'(\mu)=&\frac{I}{\rho}\frac{-V''(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{I}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)^2}H''(\mu)(\nu-\mu)\\
=&-\frac{I}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\left(V''(\mu)+\left(\frac{\rho}{I}V(\mu)+\frac{C(I)}{I}\right)H''(\mu)\right)>0\\
\implies& -V''(\mu)-\left(\frac{\rho}{I}V(\mu)+\frac{C(I)}{I}\right)H''(\mu)>0\\
\implies& -V''(\mu)-\left(\frac{\rho}{I'}V(\mu)+\frac{C(I')}{I'}\right)H''(\mu)>0,\ \forall I'>0\\
\implies& \rho V(\mu)>\sup_{I'}\left(-I'\frac{V''(\mu)}{H''(\mu)}-C(I')\right).
\end{aligned}$$

The second line uses the equality in Equation (35). The forth line uses the optimality of I in Equation (35). Note that $-I'\frac{V''(\mu)}{H''(\mu)}$ is exactly the payoff from Gaussian learning. At non-differentiable points $\mu\neq\mu^*$ of ν, $D^2V(\mu,\sigma)\leq\max\{V''(\mu-)\sigma^2,V''(\mu+)\sigma^2\}$. So the strict inequality still holds for the generalized second derivative. Therefore, Gaussian signals are strictly suboptimal. Moreover, consider:

$$\begin{aligned}
V^-(\mu)=&\max_{\nu\leq\mu,I}\frac{I}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-C(I)\\
\implies V^{-\prime}(\mu)=&-\frac{I}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\left(V''(\mu)+\left(\frac{\rho}{I}V^-(\mu)+\frac{C(I)}{I}\right)H''(\mu)\right).
\end{aligned}$$

$V^-(\mu)$ is the payoff from searching in the region $\nu\leq\mu$, given the value function $V(\mu)$. By the definition of μ^*, $V^-(\mu^*)=V(\mu^*)$. Whenever $V(\mu)=V^-(\mu)$, $V^{-\prime}(\mu)<0$. Therefore, $V^-(\mu)$ can never cross $V(\mu)$ from below — $V^-(\mu)$ is lower than $V(\mu)$.

In summary, I construct a policy function $\nu(\mu)$ and value function $V(\mu)$ solving Equation (3). Now, consider the five properties in Theorem 2. First, by construction, $\mu^*\in\arg\min V$. Second, $E=\{\mu>\mu^{**}|V(\mu)>F(\mu)\}$ is a union of disjoint open intervals $E=\bigcup I_m$. By construction, $V(\mu)=V_{\mu^m}(\mu)|_{\mu\in I_m}$. On each I_m, $\nu(\mu)$ strictly decreases and jumps down finite times. On (μ^*,μ^{**}), $\nu(\mu)$ also strictly decreases and jumps down finite times. Finally, the uniqueness argument in Lemma B.4 implies that except for those discontinuous points of ν, ν is uniquely defined. The number of such discontinuous points is countable and thus of zero measure. $p(\mu)$ is defined as $I(\mu)/J(\mu,\nu(\mu))$. □

Lemma B.1. *$I^+=\{\mu|\overline{V}^+(\mu)>F(\mu)\}$ and $I^-=\{\mu|\overline{V}^-(\mu)>F(\mu)\}$ are connected open intervals. $\{\mu|F'(\mu)>0\}\subset I^+$ and $\{\mu|F'(\mu)<0\}\subset I^-$. If $I^+\cap I^-\neq\emptyset$, $\exists!\mu^*$ s.t. $\overline{V}^+(\mu^*)=\overline{V}^-(\mu^*)>F(\mu^*)$.*

Proof. The Berge's maximum theorem implies that $\overline{V}^+$ is continuous. $\forall I>0,\nu\geq\mu$,

$$\begin{aligned}\frac{F(\nu)-\frac{C(I)}{I}J(\mu,\nu)}{1+\frac{\rho}{I}J(\mu,\nu)}\geq&\frac{F(\mu)+F'(\mu)(\nu-\mu)-\frac{C(I)}{I}J(\mu,\nu)}{1+\frac{\rho}{I}J(\mu,\nu)}\\ =&F(\mu)+\frac{F'(\mu)(\nu-\mu)-(\frac{C(I)}{I}+\frac{\rho}{I}F(\mu))J(\mu,\nu)}{1+\frac{\rho}{I}J(\mu,\nu)}\\ =&F(\mu)+F'(\mu)(\nu-\mu)-O((\nu-\mu)^2).\end{aligned}$$

Therefore, $\{\mu|F'(\mu)>0\}\subset I^+$. On any open interval that $\overline{V}^+(\mu)>F(\mu)$, the envelope theorem implies that:

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\overline{V}^+(\mu)=\frac{-H''(\mu)(\nu-\mu)(C(I)+\frac{\rho}{I}F(\nu))}{(1+\frac{\rho}{I}J(\mu,\nu))^2}>0.$$

If I^+ contains more than one interval, then at least one interval is a subset of $\{\mu|F'(\mu)\leq 0\}$, contradicting $\overline{V}^{+\prime}>0$ on the interval. A symmetric argument shows that $\{\mu|F'(\mu)<0\}\subset I^-$ and I^- is an open interval.

Suppose $I^+\cap I^-\neq\emptyset$, then on the intersection, $\overline{V}^+$ strictly increases from F and $\overline{V}^-$ strictly decreases to F. Therefore, there is a unique crossing point μ^*. □

Lemma B.2. $\inf\{\mu|U(\mu)>F(\mu)\}>0$.

Proof. Let $\mu_0=\inf\{\mu|F'(\mu)>F'(0)\}>0$. Then $\forall\mu<\mu_0$, if $U(\mu)>F(\mu)$, the optimal posterior $\nu\geq\mu_0$. By Assumption 3, $\lim_{\mu\to 0}|H'(\mu)|=\infty$, there exists δ s.t. $\forall\mu\leq\delta$, $J(\mu,\mu_0)>\frac{\sup F+\sup|F'|}{\inf F}$. Therefore, $\forall\mu\in(\delta,\mu_0)$, $U(\mu)<F(\mu)$. □

Lemma B.3. *Assume $\mu_0\geq\mu^*$, $V_0\in[F(\mu_0),\overline{V}^+(\mu_0)]$, and V_0,V_0' satisfies:*

$$V_0=\max_{\nu\geq\mu_0,I}\frac{I}{\rho}\frac{F(\nu)-V_0-V_0'(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}. \tag{40}$$

If $V_0 > F(\mu_0)$ or $U(\mu) > F(\mu)$ in a right neighbourhood of μ_0, then $\exists \mu_1 > \mu_0$ and a C^1 and strictly increasing function $V(\mu)$ defined on $[\mu_0, \mu_1)$ satisfying:

$$V(\mu) = \max_{\nu \geq \mu, I} \frac{I}{\rho} \frac{F(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \tag{41}$$

and initial conditions $V(\mu_0) = V_0$, $V'(\mu_0) = V'_0$. $V(\mu) > F(\mu)$ on (μ_0, μ_1). The maximizer $\nu(\mu)$ of Equation (41) is piecewise C^1 and strictly decreases. ν is unique in its differentiable regions. When $\mu \to \mu_1 -$, $V(\mu) \to F(\mu)$, $V'(\mu) \to F'(\mu)$.

Proof. Since $F(\mu)$ is a convex piecewise linear function, I index the actions such that $F_m = E_\mu[u(a_m, x)]$ coincides with the m-th linear piece of $F(\mu)$. Let $\underline{\mu}_m$ be the m-th kink of F: $F(\mu) = F_m(\mu) \iff \mu \in [\underline{\mu}_{m-1}, \underline{\mu}_m]$. Let $\overline{m}$ be the smallest index s.t. $F'_{\overline{m}} \geq 0$.

Step 1. Equation (40) and $V_0 \leq \overline{V}^+(\mu_0)$ implies $V'_0 \geq 0$. $\exists$ the smallest m s.t.

$$V_0 = \max_{\nu \geq \mu_0, I} \frac{I}{\rho} \frac{F_m(\nu) - V_0 - V'_0(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}. \tag{42}$$

$F'_m > 0$, because otherwise the RHS is negative. $V_0 \geq F(\mu_0)$ implies $V_0 > 0$. Suppose $V_0 = F_m(\mu_0)$, then $V_0 > 0$ implies $V'_0 < F'_m$. Then, $U(\mu_0) < V_0 = F(\mu_0)$ and this contradicts the assumption; hence $V_0 > F_m(\mu_0)$. Define:

$$\overline{V}^+_m(\mu) = \max_{\nu \geq \mu, I} \frac{I F_m(\nu) - C(I) J(\mu,\nu)}{I + \rho J(\mu,\nu)}. \tag{43}$$

Then, $V'_0 \geq 0$ and Equation (42) implies $V_0 \leq \overline{V}^+_m(\mu_0)$. Therefore, μ_0, V_0, V'_0 and m satisfies the conditions of Lemma B.4. There exists $\mu_m > \mu_0$ and strictly increasing $V_m(\mu)$ defined on $[\mu_0, \mu_m)$ satisfying

$$V_m(\mu) = \max_{\nu \geq \mu, I} \frac{I}{\rho} \frac{F_m(\nu) - V_m(\mu) - V'_m(\mu)(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}. \tag{44}$$

Let $\widehat{\nu}_m(\mu)$ be the unique maximizer of Equation (44).

Step 2. Define

$$U_{m-1}(\mu) = \sup_{\nu \geq \mu, I} \frac{I}{\rho} \frac{F_{m-1}(\nu) - V_m(\mu) - V'_m(\mu)(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}.$$

If $V_m \geq U_{m-1}$, define $V_{m-1} = V_m$ and skip to the next step. Otherwise, let $\widehat{\mu}_m = \inf\{\mu \in [\mu_0, \mu_m) | U_{m-1}(\mu) > V_m(\mu)\}$. Note that U_{m-1} is continuous in the region $V_m > F_{m-1}$ by the Berge's theorem. So $U_{m-1}(\widehat{\mu}_m) = V_m(\widehat{\mu}_m)$ when $V_m(\widehat{\mu}_m) > F_{m-1}(\widehat{\mu}_m)$. If $V_m(\widehat{\mu}_m) = F_{m-1}(\widehat{\mu}_m)$, then $U_{m-1}(\widehat{\mu}_m) = \infty$ when $V'_m(\widehat{\mu}_m) < F'_{m-1}$ and $U_{m-1}(\widehat{\mu}_m) = 0$ otherwise; both cases are impossible.[34] The definition of U_{m-1} and $V'_m > 0$ implies $V_m(\widehat{\mu}_m) < \overline{V}^+_{m-1}(\widehat{\mu}_m)$. Therefore, $\widehat{\mu}_m, V_m(\widehat{\mu}_m), V'_m(\widehat{\mu}_m)$ and m satisfies the conditions of Lemma B.4. There ex-

[34] The first case leads to strictly lower $\widehat{\mu}_m$ and the second case leads $U_{m-1} < V_m$ for an open neighbour of $\widehat{\mu}_m$.

ists $\mu_{m-1}>\widehat{\mu}_m$ and C^1 and strictly increasing V_{m-1} defined on $[\widehat{\mu}_m,\mu_{m-1})$ s.t.

$$V_{m-1}(\mu)=\max_{\nu\geq\mu,I}\frac{I}{\rho}\frac{F_{m-1}(\nu)-V_{m-1}(\mu)-V'_{m-1}(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}.$$

Extend the definition of V_{m-1} to $V_{m-1}(\mu)=V_m(\mu)$ for $\mu\in[\mu_0,\mu_{m-1})$.

Step 3. Apply the construction in *step 2* to V_{m-1} and $m-2$. Iterate until $\underline{m}$ is reached. Then, I obtain a C^1 and strictly increasing function $V_{\underline{m}}$ on $[\mu_0,\mu_{\underline{m}})$. When $\mu\to\mu_{\underline{m}}-$, $V_{\underline{m}}(\mu)\to F(\mu_{\underline{m}})$, $V'_{\underline{m}}(\mu)\to F'(\mu_{\underline{m}})$. There is a piecewise C^1 and strictly decreasing policy function $\nu(\mu)$ that $\nu(\mu_{\underline{m}})\to\mu_{\underline{m}}$.

In the construction, $\forall k$ and $\mu\geq\widehat{\mu}_k$, I only check whether $F_{k'}$ with $k'<k$ is optimal but ignored $k'>k$. Lemma S.7 implies that this is without loss of optimality:

$$V_{\underline{m}}(\mu)=\max_{\nu\geq\mu}\frac{I}{\rho}\frac{F(\nu)-V_{\underline{m}}(\mu)-V'_{\underline{m}}(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}. \tag{45}$$

The proof of Lemma S.7 uses the same technique as the derivation of Equation (35) in the proof of Theorem 2; hence, it is relegated to the supplemental materials.

Step 4. The construction does not rule out the possibility that $V_{\underline{m}}(\mu)=F(\mu)$ for $\mu\in(\mu_0,\mu_{\underline{m}})$. Suppose such μ exists, let μ_1 be the minimum of them (the equality also holds at μ_1 by continuity). Since either $V_0>F(\mu_0)$ or $U(\mu)>F(\mu)$, $\mu_1>\mu_0$. Equation (45) implies that $V'_{\underline{m}}(\mu_1)=F'(\mu_1)$. If such μ does not exist, define $\mu_1=\mu_{\underline{m}}$. Then, $V_{\underline{m}}$ pastes smoothly to F at μ_1. □

Lemma B.4. *Assume $\mu_0\geq\mu^*$, $F'_m\geq0$, $V_0\in(\max\{F_m(\mu_0),0\},\overline{V}^+_m(\mu_0)]$, and V_0,V'_0 satisfies:*

$$V_0=\max_{\nu\geq\mu_0,I}\frac{I}{\rho}\frac{F_m(\nu)-V_0-V'_0(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}. \tag{42}$$

Then $\exists\mu_1>\mu_0$ and a C^1 and strictly increasing $V(\mu)>F_m(\mu)$ defined on $[\mu_0,\mu_1)$ satisfying:

$$V(\mu)=\max_{\nu\geq\mu,I}\frac{I}{\rho}\frac{F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho} \tag{44}$$

and initial conditions $V(\mu_0)=V_0$, $V'(\mu_0)=V'_0$. The unique maximizer $\nu(\mu)$ of Equation (44) is C^1 and strictly decreases. When $\mu\to\mu_1-$, $V(\mu)\to F_m(\mu)$, $V'(\mu)\to F'_m(\mu)$ and $\nu(\mu)\to\mu_1$.

Proof. I start by deriving FOC and SOC for Equation (44):

$$\text{FOC-}\nu\text{: } \frac{I}{\rho}\left(\frac{F'_m-V'(\mu)}{J(\mu,\nu)}+\frac{F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)^2}(H'(\nu)-H'(\mu))\right)=0;$$

$$\text{FOC-}I\text{: } \frac{1}{\rho}\left(\frac{F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-C'(I)\right)=0;$$

$$\text{SOC: } \quad \text{Hess}=\begin{bmatrix}\frac{-2(H'(\mu)-H'(\nu))(\text{FOC-}\nu)}{J(\mu,\nu)}+\frac{I}{\rho}\frac{(F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu))H''(\nu)}{J(\mu,\nu)^2} & \frac{1}{I}\text{FOC-}\nu \\ \frac{1}{I}\text{FOC-}\nu & -\frac{C''(I)}{\rho}\end{bmatrix}.$$

Since $\text{Hess}_{I,I}<0$, there exists a unique I satisfying the FOC given μ,ν. On the other hand, the sign of FOC-ν is independent of I. $\text{Hess}_{\nu,\nu}<0$ when FOC-$\nu\geq 0$. Therefore, the solution of FOC-ν is unique given $I>0$ and μ. When the FOCs are satisfied, Hess is strictly negative definite. Therefore, FOC-ν and FOC-I uniquely characterize the optimal choice of ν,I. At the optimum, the equality in Equation (44) holds:

$$V(\mu)=\frac{I}{\rho}\frac{F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}-\frac{C(I)}{\rho}. \tag{46}$$

Under Equation (46), the FOCs reduce to:

$$\text{FOC-}\nu\text{:}(F_m'-V'(\mu))+\frac{\rho V(\mu)+C(I)}{I}(H'(\nu)-H'(\mu))=0; \tag{47}$$

$$\text{FOC-}I\text{:}\, IC'(I)=\rho V(\mu)+C(I). \tag{48}$$

Differentiating FOC-I implies:

$$\begin{cases} \dot{I}=\frac{\rho}{IC''(I)}(F_m'+C'(I)(H'(\nu)-H'(\mu))) \\ J(\nu,\mu)=\frac{1}{\rho}\left(I-\frac{C(I)+\rho F_m(\mu)}{C'(I)}\right) \end{cases}. \tag{49}$$

Since $J(\nu,\mu)$ is strictly monotone for $\nu\geq\mu$, define a function M implicitly by $J(M(y,\mu),\mu)=y$ to eliminate ν in the equation. Therefore, I get an ODE:

$$\dot{I}=\frac{\rho}{IC''(I)}\left(F_m'+C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I-\frac{C(I)+\rho F_m(\mu)}{C'(I)}\right),\mu\right)\right)-H'(\mu)\right)\right). \tag{50}$$

Let the RHS of Equation (50) be denoted by $R(\mu,I)$. Define $\underline{I}_m(\mu)$ implicitly by $\underline{I}_m(\mu)C'(\underline{I}_m(\mu))-C(\underline{I}_m(\mu))=\rho F_m(\mu)$ when $F_m(\mu)\geq 0$ and $\underline{I}_m(\mu)=0$ when $F_m(\mu)<0$. Since $F_m(\mu)$ increases in μ, $\underline{I}_m(\mu)$ increases (strictly increases when $\underline{I}_m(\mu)>0$). Then, $\max\{F_m(\mu_0),0\}<V_0\Longrightarrow \underline{I}_m(\mu_0)<I_0$. Define $I_m^*(\mu)$ implicitly by $I_m^*(\mu)C'(I_m^*(\mu))-C(I_m^*(\mu))=\overline{V}_m^+(\mu)$. Then, $V_0\leq\overline{V}_m^+(\mu_0)\Longrightarrow I_0\leq I_m^*(\mu_0)$. Now, I verify that Equation (50) has a solution for the initial condition $\underline{I}_m(\mu_0)<I_0\leq I_m^*(\mu_0)$.

- *Domain*: $\frac{\mathrm{d}(C'(I)I-C(I)-\rho F_m(\mu))}{\mathrm{d}I}=C''(I)I>0$ when $I>0$. When $F_m(\mu)<0$, $C'(I_m(\mu))I_m(\mu)-C(I_m(\mu))-\rho F_m(\mu)>0$. When $F_m(\mu)\geq 0$, $C'(I_m(\mu))I_m(\mu)-C(I_m(\mu))-\rho F_m(\mu)=0$. Therefore, $\forall I\geq I_m(\mu)$, $I-\frac{C(I)+\rho F_m(\mu)}{C'(I)}\geq 0$ and strict inequality holds when $I>I_m(\mu)$. Since M only applies to non-negative reals, Equation (50) is well defined on $\{I|I\geq\underline{I}_m(\mu)\}$.

- *Continuity*: When μ is strictly bounded away from $\{0,1\}$, and I is uniformly bounded away from $\underline{I}_m(\mu)$, the conditions for the Picard-Lindelof theorem hold.[35]

[35] It is sufficient to verify that $\frac{\partial}{\partial y}M(y,\mu)=-\frac{1}{H''(M(y,\mu))(M(y,\mu)-\mu)}$ is bounded. $M(y,\mu)=\mu$ implies $J(\nu,\mu)=0$ and $\frac{1}{\rho}\left(I-\frac{C(I)+\rho F_m(\mu)}{C'(I)}\right)=0$. Since I is uniformly bounded away from $\underline{I}_m(\mu)$, then $M(y,\mu)-\mu$ is uniformly bounded away from 0.

- *Monotonicity*: When $I = I_m^*(\mu)$, the optimality condition of $\overline{V}_m^+$ implies:

$$\begin{cases} F_m' - C'(I)(H'(\mu) - H'(\nu)) = 0 \\ (I + \rho J(\mu,\nu))C'(I) = C(I) + \rho F_m(\nu) \end{cases} \Longrightarrow (I - \rho J(\nu,\mu))C'(I) = C(I) + \rho F_m(\mu)$$
$$\Longrightarrow F_m' + C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I - \frac{C(I)+\rho F_m(\mu)}{C'(I)}\right),\mu\right)\right) - H'(\mu)\right) = 0$$
$$\Longrightarrow \dot{I} = R(\mu, I) = 0.$$

Then, consider the monotonicity of $R(\mu, I)$:

$$\frac{\partial}{\partial I}R(\mu,I) = C''(I)\big(H'(M) - H'(\mu)\big) + C'(I)\frac{H''(M)}{H''(M)(\mu - M)}\frac{1}{\rho}\frac{C(I)+\rho F_m(\mu)}{C'(I)^2}C''(I) < 0.$$

Therefore, $R(\mu, I)$ is positive in the region $\{\underline{I}_m(\mu) < I < I_m^*(\mu)\}$.

$\forall \delta > 0$, The Picard-Lindelof theorem guarantees a unique solution $I(\mu)$ satisfying the ODE in the region $\mu \in [\delta, 1-\delta], I \in [\underline{I}_m(\mu) + \delta, I_m^*(\mu)]$. $I(\mu)$ lies between $\underline{I}_m(\mu)$ and $I_m^*(\mu)$ and increases until it hits $\underline{I}_m(\mu) + \delta$. Next, let $\delta \to 0$ and extend $I(\mu)$ toward the boundary. By the monotonicity of $\underline{I}_m$, there exist limits $\bar{I}, \bar{\mu}$ satisfying $\underline{I}_m(\bar{\mu}) = \bar{I}$. Then, since $R(\mu, I)$ has a limit $\frac{\rho F_m'}{\bar{I}h''(\bar{I})}$, $\lim_{\mu \to \bar{\mu}} V'(\mu) = F_m'$ by Equation (47). So the resulting $V(\mu)$ smoothly pastes to F_m at $\bar{\mu}$. Equation (48) implies that $V(\mu)' > 0$ on $[\mu_0, \bar{\mu})$.

Since M is a C^1 function, the optimal posterior $\nu(\mu) = M(\frac{1}{\rho}(I(\mu) - \frac{C(I(\mu))+\rho F_m(\mu)}{C'(I(\mu))}), \mu)$ is C^1 on $[\mu_0, \bar{\mu})$. Finally, differentiating Equations (46) and (47) yields:

$$\begin{cases} V'(\mu) = \frac{I(\mu)}{\rho}\frac{\nu(\mu)-\mu}{J(\nu(\mu),\mu)}(-V''(\mu) - C'(I(\mu))H''(\mu)) > 0 \\ -V''(\mu) + C''(I(\mu))I'(\mu)(H'(\nu(\mu)) - H'(\mu)) + C'(I(\mu))(H''(\nu(\mu))\nu'(\mu) - H''(\mu)) = 0 \end{cases}$$
$$\Longrightarrow \nu'(\mu) = \frac{(V''(\mu) + C'(I(\mu))H''(\mu))}{J(\nu(\mu),\mu)C'(I(\mu))H''(\nu(\mu))}(H(\mu) - H(\nu(\mu)) + H'(\mu)(\nu(\mu) - \mu))$$
$$- \frac{\rho V'(\mu)}{I(\mu)C'(I(\mu))H''(\nu(\mu))}(H'(\nu(\mu)) - H'(\mu))$$
$$= -\frac{\rho V'(\mu)}{I(\mu)(\nu(\mu) - \mu)C'(I(\mu))H''(\nu(\mu))}(H(\mu) - H(\nu(\mu)) - H'(\nu(\mu))(\mu - \nu(\mu))) < 0.$$

The first equality is from $I'(\mu)C''(I(\mu)) = \frac{\rho V'(\mu)}{I(\mu)}$ (Equations (47), (48) and (49)). The second equality is from the equation for V'. Therefore, $\nu(\mu)$ is C^1 and strictly decreases on $[\mu_0, \bar{\mu})$. When $\mu \to \bar{\mu}$, since $V(\mu) \to F_m(\mu)$, the maximizer $\nu(\mu) \to \bar{\mu}$. □